Working Paper

Document de travail

Fiscal Policy and
Economic Analysis
Branch

Department of Finance

Direction de la
politique fiscale
et de l'analyse économique

Ministère des Finances

The Constant Price Hypothesis in Multi-Industry Cross-Section Studies: The Impact on Cost Function Estimates

by

Benoit Robidoux
and
John Lester

Working Paper No. 90-04

Economic Studies and Policy Analysis Division
Fiscal Policy and Economic Analysis Branch
Department of Finance, Canada

The data used in this study are from Statistics Canada's Census of Manufactures. Census data for individual establishments were accessed through the Business and Labour Market Analysis Group. The authors wish to thank Richard Hotte and Brenda Hutchinson for providing expert programming assistance. They are also indebted to Michael Smart, Pierre Mohnen and Janice McMechan for useful suggestions and comments. Any errors remain the sole responsibility of the authors.

Working papers are intented to make analytic work undertaken in the Department of Finance available to a wider readership. They have received only limited evaluation and do not necessarily reflect the views of the Department of Finance. Comments on the working papers are invited and may be sent to the author(s). They are circulated in the language of preparation.

ABSTRACT

Econometric estimation of cost functions with plant data is often used to evaluate scale economies in multi-industry cross-section studies. However, unlike time-series analysis, multi-industry micro studies generally assume that input and output prices are constant across plants in a given industry. This approach could obviously produce biased results if input and output prices vary across plants. This study tests this hypothesis and its impact on cost function estimates by calculating plant-level Divisia-Wald price indexes for output and inputs and using them in estimating unit cost functions for a sub-sample of twenty Canadian manufacturing industries in 1979. In addition, the analysis considers explicitly differences in production structure among Canadian regions by including regional dummies in the cost function.

The results suggest that input and output prices vary among plants of a given industry. The results also indicate that explicitly taking into account the output price variance substantially increases, on average, estimates of economies of scale and region-specific cost differences. In contrast, the inclusion of intermediate input prices reduces marginally, on average, scale economies and region-specific cost differences. However, the output price effect generally dominates and, consequently, the results suggest that scale economies and region-specific cost differences derived from cost functions under the standard constant price assumption are underestimated on average. Finally, even if this average picture holds in most industries, some industry results are not affected or are affected in the opposite direction when the constant price hypothesis is relaxed.

RÉSUMÉ

L'estimation économétrique de fonctions de coût à partir de données en coupe transversale est une méthode fréquemment utilisée pour évaluer les économies d'échelle internes à l'établissement. Cependant, au contraire des analyses utilisant des séries choronologiques, les études micro-économiques supposent généralement que les prix des inputs et de l'output sont constants pour tous les établissments dans une même industrie. Or, les résultats découlant de ces études peuvent être biaisés substantiellement si cette hypothèse n'est pas respectée. Cette étude teste cette hypothèse et son impact sur l'estimation de fonctions de coût en calculant des indices de prix Divisia-Wald au niveau de l'établissment pour les inputs et l'output et en les utilisant dans l'estimation de fonctions de coût pour un sous-échantillon de vingt industries manufacturières canadiennes en 1979. De plus, l'analyse prend explicitement en compte les différences de structure de production entre les régions canadiennes en incluant des variables muettes régionales dans la fonction de coût.

Les résultats suggèrent que les prix des inputs et de l'output varient entre les établissement d'une même industrie. Les résultats indiquent également que la prise en compte explicite de la variance du prix de l'output augmente, en moyenne, les estimations d'économies d'échelle et de différences de coût spécifiquement régionales. Inversement, l'inclusion du prix des intrants intermédiaires réduit marginalement, en moyenne, les économies d'échelle et les différences de coût spécifiquement régionales. En bout de ligne, cependant, l'effet provenant du prix de l'output domine largement et, par conséquent, les résultats suggèrent que les économies d'échelle et les différences de coût spécifiquement régionales dérivées de fonctions de coût estimées sous l'hypothèse usuelle de prix constants sont sous-estimées en moyenne. Finalement, même si cette conclusion générale tient dans la plupart des industries analysées, les résultats ne sont pas affectés ou le sont inversement dans certaines industries, lorsqu'on relâche l'hypothèse de prix constant.

TABLE OF CONTENTS

		Page
I.	Introduction	1
II.	Empirical Specification	2
III.	The Divisia-Wald Price Indexes	9
IV.	Data	12
V.	Variance and Correlation Analysis	13
VI.	Regression Results	17
VII.	Conclusion	24
	References	26
	Annex	26
	A1 Construction of Price Indexes and Selection of the Final Sample	28
	A2 Representativity of the Final Sample	32

LIST OF TABLES

Table 1 Mean and Standard Deviation of Price Indexes by Industry

Table 2 Pearson Coefficients of Correlation for all the Industry Sample After Scaling Relevant Variables

Table 3 Cost Disadvantage Ratio (CDR) at 1/2 MES

Table 4 Cost Saving Achievable by Eliminating Scale-Related Inefficiency by Region Percentage

Table 5 Cost Saving Achievable by Eliminating Region-Related Inefficiency by Region Percentage

Table 6 Ratio of Full to Nominal Aggregate Estimates

Table A1 Selection of the Final Regression Sample: Preliminary Regression Results and Data Information

Table A2 Gross Output Share by Region in the Regression Sample

Table A3 Comparing the Census Sample and the Regression Sample: Distribution in Percentage of Key Industry Variables by Region Relative to the Canadian Level

Table A4 Estimated Coefficients of Equation 1 (Nominal Estimates)

Table A5 Estimated Coefficients of Equation 2 (Real Estimates)

Table A6 Estimated Coefficients of Equation 3 (Full Estimates)

Table A7 Differences Between Full and Nominal Estimates: Hausman Tests

I. INTRODUCTION

Econometric estimation of cost or production functions with plant data is a method often used to evaluate scale economies in multi-industry cross-section studies.[1] However, unlike macro studies, multi-industry micro studies generally assume that input and output prices are constant across plants in a given industry. Gupta (1968), Griliches and Ringstad (1971), De Brabander and Vanlommel (1978), Corbo and Meller (1979), Fuss and Gupta (1981) and Baldwin and Gorecki (1986) are examples of micro studies implicitly using this assumption. This hypothesis could create errors-in-variables and specification biases if input and output prices vary across plants. In general, the size and the direction of these biases cannot be assessed without empirical analysis.

In the Canadian manufacturing case, Robidoux and Lester (1986) found that assuming constant wage rates induces, on average, a downward bias on economies of scale, reflecting a positive correlation between wage rates and output in the sample used.[2] However, this bias was present only in some industries: in most industries, wage differences compensated for productivity differences with the result that production costs are unaffected. In this study, we pursue this work further by calculating plant-level Divisia-Wald price indexes for gross output, materials and energy and using them in estimating a unit cost function for a sub-sample of twenty four-digit Canadian manufacturing industries in 1979. It is then possible to evaluate how the standard assumption of constant output and input prices affects the results of cross-section cost studies. In addition, the analysis considers explicitly differences in production structure among Canadian regions by including regional dummy variables in the cost function. This procedure provides results on region-specific cost inefficiencies which are interesting, per se, in the Canadian context.

---

1. Engineering and survivor methods are other approaches less popular among economists.

2. Note, however, that no formal statistical tests were carried out by Robidoux and Lester (1986). Their results are based on point estimate comparisons.

The results obtained suggest that assuming constant output prices among plants creates a substantial negative bias, on average, on scale economies and on region-specific cost inefficiencies in all Canadian regions. On the other hand, taking into account materials and energy prices marginally reduces, on average, scale economies and region-specific inefficiencies. However, the output price effect generally dominates; consequently scale economies and region-specific inefficiencies derived from cost or production functions under the standard constant price assumption are underestimated. Not surprisingly in some industries estimates of scale economies and regional inefficiencies are unaffected and in others overstated by the assumption of constant prices.

This paper is divided in six sections and contains two annexes. The empirical specification of the cost function and estimation procedure are presented in Section II. The Divisia-Wald price index used for output, materials and energy is described in Section III. Section IV provides preliminary analysis of micro data used in the regression analysis and the results are presented in Section V. The final section contains concluding remarks. Annexes provide details about the construction of price indexes and describe the selection and the representativeness of the manufacturing sample used.

II. EMPIRICAL SPECIFICATION

One important characteristic of our data base -- the Census of Manufactures -- should be taken into account in specifying the unit cost function: the lack of plant-specific data for capital.[3] As a result of this lack of capital data the cost function specification used in this study is based on the assumption that the production process is non-homogeneous but

---

3. In fact, no capital data at the plant level are available in the Census of Manufactures or elsewhere since the investment survey is done at the firm level.

homothetic.[4] Non-homogeneous cost functions have the desirable property of not constraining the scale elasticity to be the same at all output levels, as is the case for homogeneous functional forms such as the well known Cobb-Douglas. The homotheticity assumption means that factor proportions depend only on relative prices, not on output levels. Moreover, as shown by Fuss and Gupta (1981), this property of homothetic functions permits exclusion of capital costs (or other factor costs) from the left-hand side of the unit cost equation without affecting the scale economy parameters.

In line with Fuss and Gupta (1981), the specific homothetic form used is the semi-logarithmic function. In order to evaluate the impact of excluding output and input prices, three equations are estimated:[5]

$$(1)\quad \ln(C/V) = \alpha_1 + \alpha_2 V + \alpha_3(1/V) + \alpha_4 W1 + \alpha_5 W2 + \alpha_6 DIV + \alpha_7 D1 + \alpha_8 D2 + \alpha_9 D3 + \alpha_{10} D4 + \mu \text{ (nominal form)}$$

$$(2)\quad \ln(C/Q) = \alpha_1 + \alpha_2 Q + \alpha_3(1/Q) + \alpha_4 W1 + \alpha_5 W2 + \alpha_6 DIV + \alpha_7 D1 \quad \alpha_8 D2 + \alpha_9 D3 + \alpha_{10} D4 + \mu \text{ (real form)}$$

$$(3)\quad \ln(C/Q) = \alpha_1 + \alpha_2 Q + \alpha_3(1/Q) + \alpha_4 W1 + \alpha_5 W2 + \alpha_6 DIV + \alpha_7 D1 + \alpha_8 D2 + \alpha_9 D3 + \alpha_{10} D4 + \alpha_{11} P_m + \alpha_{12} P_e + \mu \text{ (full form)}$$

where: C is operating or non-capital costs; V is the value of gross output; Q is the gross output quantity index; W1 and W2 are the wage rates of production

---

4. Although there is little likelihood that this hypothesis is strictly correct, empirical evidence surveyed in Robidoux (1987) suggests it is a tolerable approximation when it is used to exclude capital costs.

5. These specific cost functions could be derived from the general homothetic functional form, $C(P,Q) = h(Q)\, g(P)$ where P is a vector of input prices and other variables affecting costs.

and non-production workers, respectively; DIV is a product diversity index; D1, D2, D3 and D4 are dummy variables equal to one when the plant is in Atlantic provinces, Quebec, Prairies provinces and British Columbia, respectively, and zero otherwise, the Ontario dummy being excluded; P_q, P_m and P_e are the Divisia-Wald price indexes for gross output, materials and energy, respectively.[6]

In the first equation, output is measured in value terms and intermediate input prices are not taken into account. Although in many studies wages rates, a product diversity index or regional dummies are not included, this equation corresponds to the usual cost function estimated in the sense that constant input and output prices are assumed across plants. In contrast, the second equation relaxes the assumption of constant output price and the third equation adds materials and energy prices in the right-hand side of the unit cost equation.[7]

The constant price hypothesis could create different types of bias on the estimated parameters.[8] First, assuming a constant output price when in fact this price varies among plants could be interpreted as an errors-in-variables bias affecting both dependent and independent variables. The standard econometric analysis says that errors in measuring the dependent variable are incorporated in the disturbance term and cause no problems. In contrast, if there are errors in measuring independent variables, these variables become stochastic and parameter estimates are biased, even

---

6. The wage rate of non-production workers is calculated as the wage bill of administrative, office and other non-manufacturing employees divided by usual hours worked. Usual hours worked correspond to standard weekly hours times the number of weeks effectively worked (i.e., 52 less paid vacation time and statutory holidays, both expressed in weeks). The product diversity index is the reciprocal of the Herfindahl index and varies between 1 and infinity, an higher value indicating a higher level of diversity.

7. However, it is worthwhile to note that the capital price is still assumed to be the same for all plants in a given industry in the three equations.

8. A formal analysis of what follows in this paragraph could be found in any standard econometric textbook.

asymptotically. If only one independent variable is measured with error, it can be shown that the bias on the coefficient of this variable is unambiguously negative. The coefficients of the other variables are also biased but the direction of bias can be either way. However, all these results are based on the assumptions, inter alia, that errors in measuring dependent and independent variables are both mutually uncorrelated and uncorrelated with the correctly measured variables. In our case, these assumptions are probably not valid since we could expect a non-zero correlation between output volume, output price and real unit cost.[9] If we relax these assumptions, the analysis become more complicated. In this case even errors in measuring the dependent variable create biases and the direction of these biases on independent variables (measured with and without errors) is ambiguous without knowledge of the sign and the size of many variances and covariances (see Maddala (1977, p. 302)).

Second, omitting the materials and energy prices from the independent variables creates standard specification biases. The direction of these biases on the other independent variables depends, as is well-known, on the correlation between these variables and the omitted price variables. In sum, we know that many types of bias potentially affect the estimation of the standard unit cost function but the direction of these biases cannot be assessed without empirical investigation.

Regional differences are a well-known characteristic of the Canadian economy. Regional dummies have been included in the cost function in order to capture systematic differences between regions other than plant size, product diversity and input prices. In specifying equations (1) to (3), we also assume that in each region the logarithmic of the unit cost function has a different intercept but the same slope, since no interaction terms between regional dummies and output appear.[10]

---

9. In fact, empirical results presented below indicate significant correlations between these variables.

10. In other terms, $\partial \ln AC_i / \partial Z_i = \alpha_2 - \alpha_3 / Z_i^2$, where $Z_i = V_i$ or Q_i and $AC_i = C_i / Z_i$, and is then the same for all regions. Note however that (continued...)

Given results for the unit cost equations, classical indicators of scale economies, such as the minimum efficient scale (MES) and the cost-disadvantage-ratio (CDR), are derived. The MES is defined as the square root of $(\hat{\alpha}_3/\hat{\alpha}_2)$ if the unit cost curve is U-shaped and as the average size of the largest plants accounting for 50 per cent of industry output (TOP-50 measure) for the other cost curve shapes exhibiting increasing returns to scale -- hyperbolic, linear and quasi-linear forms.[11] On the other hand, the CDR at one-half of MES is defined as:

$$(4) \quad CDR = \frac{\exp(\hat{\alpha}_2(MES/2) + \hat{\alpha}_3(2/MES)) - \exp(\hat{\alpha}_2 MES + \hat{\alpha}_3(1/MES))}{\exp(\hat{\alpha}_2 MES + \hat{\alpha}_3(1/MES))}$$

where $\hat{\alpha}_i$ are estimated coefficients. As indicated by equation (4), the CDR is the percentage that costs increase in going from MES to 1/2 MES. It is also worthwhile to note again that, given the specification used for the cost function, the MES and the CDR are the same for all regions.[12]

However, since average plant size varies across regions, the cost penalty associated with suboptimal capacity varies also from one region to another. In order to describe this regional dispersion in scale inefficiency, we have calculated a summary statistic called the potential scale-related cost saving achievable (srCSA) defined for region i as:

10.(...continued)
for a given output level, $\partial AC_i/\partial Z_i = [\exp(\alpha_1 + \alpha_2 Z_i + \alpha_3/Z_i + \alpha_4 W1_i + \alpha_5 W2_i + \alpha_6 DIV_i + \alpha_7 D1_i + \alpha_8 D2_i + \alpha_9 D3_i + \alpha_{10} D4_i + \alpha_{11} P_{mi} + \alpha_{12} P_{ei})](\alpha_2 - \alpha_3/Z_i^2)$ and then differs potentially between regions.

11. See Robidoux and Lester (1986) for more details on the relation between the shape of the unit cost curve and the signs of estimated coefficients.

12. It is also the case for the cost elasticity which is given by, $\varepsilon_{cz} = \alpha_2 Z - \alpha_3/Z + 1$.

$$(5) \quad srCSA_i = (\overline{AC}_i - ACMES_i)/\overline{AC}_i$$

where,

$$(6) \quad ACMES_i = \exp[\hat{\alpha}_1 + \hat{\alpha}_2 MES + \hat{\alpha}_3(1/MES) + \hat{\alpha}_4 \overline{W1}_i + \hat{\alpha}_5 \overline{W2} + \hat{\alpha}_6 \overline{DIV}_i + \hat{\alpha}_7 D1_i + \hat{\alpha}_8 D2_i + \hat{\alpha}_9 D3_i + \hat{\alpha}_{10} D4_i + \hat{\alpha}_{11} \overline{P}_{mi} + \alpha_{12} \overline{P}_{ei}]$$

and $\overline{AC}_i$ is the industry average cost in the regression sample for region i, $ACMES_i$ is predicted average cost at MES in region i, $\hat{\alpha}_i$ are estimated coefficients and a bar over a variable indicates an industry average calculated from the regression sample.[13] The $srCSA_i$ corresponds then to the potential cost saving achievable if all plants were the minimum efficient size in region i. In contrast to the CDR, which is an approximation of the steepness of the unit cost curve between 1/2 MES and MES, the srCSA has the advantage of taking into account both the slope of the unit cost curve and the size distribution of plants.[14]

The estimation of unit cost functions also permits the identification of cost differences between regions that come from region-specific characteristics such as technology, management and other factors not

---

13. In using results from equations (1) and (2), $\hat{\alpha}_{11}$ and $\hat{\alpha}_{12}$ are set to zero by definition.

14. Alternatively to $\overline{AC}_i$, the predicted average cost at the regional industry average output could be used in calculating $srCSA_i$. In general, the industry average costs are likely to be under rather than on the estimated cost curve since they correspond to a weighted average of plant unit costs where larger plants have larger weights. This will be the case if unit costs of all plants are on the estimated curve or if they are distributed symmetrically around it. In general, the data do not meet these conditions, so industry unit cost could be higher than the predicted value or even lower than $ACMES_i$. In the empirical part of this study, $scCSA_i$ was constrained to be non-negative.

included as regressors and correlated with region. Similarly to regional scale inefficiencies, assuming that the region j is the cost efficient region based on dummy coefficient estimates, the potential region-specific cost saving achievable (rsCSA) in region i is defined as:

$$(7) \quad rsCSA_i = (AC1_i - AC2_i)/\overline{AC}_i$$

where,

$$(8) \quad AC1_i = \exp(\hat{\alpha}_1 + \hat{\alpha}_2\bar{Z}_i + \hat{\alpha}_3(1/\bar{Z}_i) + \hat{\alpha}_4\overline{W1}_i + \hat{\alpha}_5\overline{W2}_i + \hat{\alpha}_6\overline{DIV}_i + \hat{\alpha}_7 D1_i + \hat{\alpha}_8 D2_i + \hat{\alpha}_9 D3_i + \hat{\alpha}_{10} D4_i + \hat{\alpha}_{11}\overline{P_{mi}} + \alpha_{12}\overline{P_{ei}})$$

$$(9) \quad AC2_i = \exp(\hat{\alpha}_1 + \hat{\alpha}_2\bar{Z}_i + \hat{\alpha}_3(1/\bar{Z}_i) + \hat{\alpha}_4\overline{W1}_i + \hat{\alpha}_5\overline{W2}_i + \hat{\alpha}_6\overline{DIV}_i + \hat{\alpha}_7 D1_j + \hat{\alpha}_8 D2_j + \hat{\alpha}_9 D3_j + \hat{\alpha}_{10} D4_j + \hat{\alpha}_{11}\overline{P_{mi}} + \hat{\alpha}_{12}\overline{P_{ei}})$$

where $\bar{Z}_i = \bar{V}_i$ or $\bar{Q}_i$. Then $rsCSA_i$ corresponds to the impact of eliminating region-specific inefficiencies on average costs of the regional industry i keeping the size distribution of plants and input prices constant.

Since the 1970s, the cost share approach is become the usual or dominant estimation procedure used in cost function studies. Under the assumption that product and input markets are competitive, the cost share equations implied by the Shephard's lemma are estimated in conjunction with the cost function using the seemingly unrelated equations (SURE) estimator. The derivation of the cost share equations are also based on the usual assumptions that each observation represents an equilibrium point and that inputs are measured in efficient units. As noted by Corbo and Meller (1979), as long as these assumptions are valid there is no difference between this

approach and a direct estimation procedure.[15] In contrast, if these assumptions are violated, it is impossible to know if the parameter estimates are those of the underlying cost function or spurious estimates resulting from a misspecification. Taking into account the potential heterogeneity of labour in a micro data context, cost functions were estimated directly.[16]

Since we use cross-section data in this study, the problem of heteroscedasticity could arise in estimation. The presence of heteroscedasticity was tested with the procedure proposed by White (1980). Based on this test, we reject the null hypothesis of homoscedasticity at 1 per cent level in all industries examined. The presence of heteroscedasticity in the disturbances leads to consistent but inefficient parameter estimates and inconsistent variance matrix estimates. In our case, we choose to deal with the latter problem by using the heteroscedasticity-consistent variance matrix proposed by White (1980). Thus, even if heteroscedasticity cannot be eliminated in terms of efficiency loss, correct inferences can be drawn.

III. THE DIVISIA-WALD PRICE INDEXES

This section presents a Divisia-Wald aggregation procedure used for calculating establishment-specific prices for energy (P_e), materials (P_m) and gross output (P_q) that are used directly or indirectly in equations (2) and (3). The problem is to compute the appropriate price index with data on quantity and value available from the Census of Manufactures. A common practice in empirical economics is to use the approximate Divisia aggregation

---

15. However, on statistical grounds, the SURE estimator is more efficient than the OLS estimator if error terms are correlated *and* if the set of explanatory variables is not the same in each equation.

16. See Robidoux and Lester (1988) for empirical evidence.

procedure.[17] An inter-establishment approximation to the Divisia price index is given by:[18]

$$(10)\quad P^j = \prod_{i=1}^{n} (p_i^j/p_i^o)^{W_i^j} \qquad \text{where,} \qquad W_i^j = 1/2(S_i^o + S_i^j)$$

and where, P^j is the appropriate (P_e, P_m or P_q) approximation of the Divisia price index of the j th establishment; p_i^j (q_i^j) is the price (quantity) of the i th type of input used or output produced in the j th establishment (i=1,.....,n and j=1,.....k); p_i^o (q_i^o) is the price (quantity) of the i the type of input used or output produced in the base establishment; $S_i^j = p_i^j q_i^j / \Sigma_i p_i^j q_i^j$ and $S_i^o = p_i^o q_i^o / \Sigma_i p_i^o q_i^o$ are the value shares of the i th type of input or output in respectively the j th establishment and the base establishment.[19] In other terms, the approximate Divisia price index corresponds to a weighted geometric mean of individual-to-base price ratios where the weights are arithmetic means of value shares in current and base establishments.

However, two decisions must be made before this index procedure becomes operational in our specific context. First, we have to define the arbitrary base establishment. It is possible to choose one of the establishments in an industry or to use an average of all the establishment data in an industry in order to represent the base establishment. Following Caves, Christensen and Diewert (1982), the second approach is used and prices and shares of the base establishment are defined as:

---

17. See Fuss(1977) and papers quoted therein for applications of the approximate Divisia aggregator.

18. This discrete approximation to the Divisia price index is in fact the Tornquist (1936) index.

19. See Hulten (1973) and Diewert (1976) for the desirable properties of this index number.

$$(11)\quad p_i^o = \prod_{j=1}^{m} (p_i^j)^{1/m} \quad \text{and} \quad S_i^o = \frac{1}{k} \left(\sum_{j=1}^{k} S_i^j \right)$$

where $m \leq k$ and corresponds to the number of plants using the i th input or producing the i th output. By this procedure, we generate a base establishment, for each industry, corresponding to a hypothetical average establishment whose prices and shares represent respectively geometric averages of prices and arithmetic averages of shares for the whole industry.[20] Caves, Christensen and Diewert (1982) have established that this index number satisfies the transitivity test and maintains a high degree of characteristicity.[21]

Second, the fact that one or more q_i^j may be equal to zero will cause some problems. This will occur since not all plants in a given industry will use exactly the same set of inputs or will produce the same type of products. If one or more q_i^j are not present in a particular plant and, consequently, the corresponding p_i^j are undefined, then the approximate Divisia price index will be also undefined.[22] This problem comes basically from the weighting form of the Divisia index which associates a non-zero weight to a component not used or produced in a given plant. Rather than using an ad hoc correction such as using the industry average as a shadow price, we have

---

20. With this procedure we have one base plant for each four-digit industry. It is also possible to define a unique base plant corresponding to an average of all the manufacturing sector data. Although this alternative approach seems to be superior for inter-industry price comparisons, the industry-specific base plant used here seems to be more appropriate for the industry-specific regression analysis.

21. Transitivity and characteristicity are two of the Fisher's (1922) criteria for choosing among index numbers. An index number, I, satisfies the transitivity test if $I_i^j = I_k^j / I_k^i$. Characteristicity indicates the degree to which weights used are specific to the components used in the index. For more details, see Allen (1975).

22. By construction, $p_i = p_i q_i / q_i$ (where q_i and $p_i q_i$ only are observed) when $q_i \neq 0$. If $q_i = 0$ then p_i is undefined or unknown.

changed the weighting procedure used. Following Wald (1901), we have used the following alternative weights:[23]

$$(12)\quad \dot{W}_i^j = (S_i^o S_i^j)^{1/2} \Big/ \sum_{i=1}^{n} (S_i^o S_i^j)^{1/2}$$

This modification implies that the i th weight vanishes when the i th price vanishes, thereby avoiding the problem created by the unknown prices. As a result, the Divisia-Wald index used to calculate plant-specific prices for energy, materials and output corresponds to:[24]

$$(14)\quad P^j = \prod_{i=1}^{n} (p_i^j / p_i^o)^{\dot{W}_i^j}$$

where all components are defined by equations (11) and (12). Basically, as for the Divisia index, this price index will capture price differences for a given product, not price differences for different products.[25]

IV. DATA

The data used to estimate the Wald-Divisia price indexes and the unit cost function are from the Census of Manufactures. Confidential census data by individual establishments have been accessed though Statistics Canada's Business and Labour Market Analysis Group. The Census collects data on manufacturing and other activities of manufacturers, such as wholesaling.

---

23. See Theil (1973) for a presentation of the Wald's approach.

24. Theil (1973) has demonstrated that weights $\dot{W}_i^j$ satisfy Fisher's factor reversal test, which requires that the sum of the log-changes in the price and quantity indexes be equal to the log-change in total value, more closely than the weights W_i^j.

25. On practical grounds this requires that each product be measured in homogeneous units. See Annex for more details.

Although we would prefer to use data based solely on manufacturing activity, the consistent set of data necessary for this study could only be created on a total activity basis.

After controlling for data quality and availability, it was possible to estimate unit cost functions represented by equations (1) to (3) in only 20 four-digit manufacturing industries. The process by which we came to this sample is presented in Annex 1, as well as details on the empirical implementation of the Divisia-Wald price indexes.

Using Census data, the 20 industries selected in this study correspond to almost one-third of total manufacturing production in 1979. On a regional basis, this proportion is higher in all regions except Ontario, where the coverage is limited to 20 per cent of the manufacturing output. In general, the sample is regionally representative of overall manufacturing in terms of unit cost and average gross output by plant except in Ontario and B.C. where plant size tends to be smaller and larger than in all manufacturing, respectively. More details are presented in Annex 2.

V. VARIANCE AND CORRELATION ANALYSIS

All the preceding analysis is based on the presumption that some input and output price variance exists in a given industry. The standard dispersion measures presented in Table 1 indicate that this is effectively the case in virtually all industries. The standard deviations of output, materials and energy price indexes are on average around 0.2 and vary between 0.04 and 0.30 with most cases over 0.10. This dispersion of input and output price indexes is similar to that for unit cost, wage rates or product diversity. It appears that, in general, plants in a given industry do not pay the same price for their inputs and do not sell their products at the same price.

TABLE 1

Mean and Standard Deviation of Price Indexes by Industry

Industry		Output Price		Materials Price		Energy Price	
		Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
1011	Slaughtering and Meat	1.02	0.26	1.01	0.18	1.02	0.16
1012	Poultry Processors	1.01	0.12	0.99	0.09	1.00	0.07
1020	Fish Products	1.07	0.18	0.94	0.29	1.09	0.17
1031	Fruits & Vegetables	1.03	0.11	0.99	0.14	1.02	0.15
1040	Dairy Products	0.98	0.12	0.97	0.13	1.00	0.11
1050	Flour & Cereal	1.04	0.14	1.02	0.15	0.97	0.17
1060	Feed	1.03	0.12	1.02	0.10	0.99	0.14
1071	Biscuits	1.03	0.23	0.98	0.08	1.04	0.21
1072	Bakeries	1.00	0.21	1.01	0.13	1.01	0.11
1089	Misc. Food Processors NES	1.03	0.21	0.98	0.16	1.00	0.18
1093	Breweries	0.99	0.14	1.00	0.11	1.02	0.20
1860	Carpet, Mat and Rug	1.09	0.28	1.05	0.13	0.97	0.08
2310	Hosiery	1.00	0.23	1.05	0.29	1.01	0.05
2392	Other Knitting	1.00	0.31	0.99	0.25	1.03	0.18
2513	Sawmills and Planing Mills	1.05	0.17	1.04	0.30	1.02	0.18
2520	Veneer and Plywood Mills	1.05	0.21	1.00	0.23	1.01	0.17
2710	Pulp and Paper Mills	1.01	0.11	1.00	0.18	1.00	0.19
2731	Corrugated Box	1.01	0.16	1.01	0.08	1.01	0.15
3550	Ready-Mix Concrete	0.99	0.18	0.99	0.18	0.99	0.14
3651	Petroleum Refining	1.00	0.10	1.02	0.04	1.02	0.32
ALL[1]		1.00	0.18	1.00	0.21	1.00	0.16

1. Average prices are normalized on one in each industry.

Before presenting regression results, it is also very instructive to look at the simple correlation coefficients between variables used in the regression analysis. These correlation coefficients are presented in Table 2 for the whole sample of 20 industries which includes 1327 plant observations. All means of non-dummy variables have been normalized on one in each industry in order to purge the variance in the mean of a given variable among different industries.[26]

26. Obviously, this analysis will capture only correlations that are relatively systematic at the industry level; marginal industry correlation and non-systematic relations will not be detected.

TABLE 2
Pearson Coefficients of Correlation for all the Industry Sample After Scaling all Relevant Variables[1]

	C/V	Q	V	Pq	Pm	Pe	W1	W2	DIV	D1	D2	D3	D4	D5
C/Q	0.5398 (0.0001)	-0.1593 (0.0001)	-0.0781 (0.0044)	0.7448 (0.0001)	0.2842 (0.0001)	0.0625 (0.0228)	0.0156 (0.5694)	-0.0068 (0.8042)	0.0017 (0.9513)	0.0368 (0.1803)	-0.0266 (0.3332)	-0.0264 (0.3361)	0.0039 (0.8860)	0.0162 (0.5560)
C/V		-0.0860 (0.0017)	-0.1048 (0.0001)	-0.1487 (0.0001)	0.1059 (0.0001)	0.0402 (0.1433)	-0.0017 (0.9519)	-0.0131 (0.6330)	0.0126 (0.6463)	0.0486 (0.0767)	-0.0041 (0.8824)	0.0202 (0.4628)	-0.0020 (0.9434)	-0.0437 (0.1118)
Q			0.9843 (0.0001)	-0.1265 (0.0001)	0.0618 (0.0244)	-0.2072 (0.0001)	0.4138 (0.0001)	0.1475 (0.0001)	0.1775 (0.0001)	-0.0718 (0.0089)	-0.0850 (0.0019)	-0.0003 (0.9903)	0.2033 (0.0001)	-0.0248 (0.3663)
V				-0.01467 (0.5935)	0.1079 (0.0001)	-0.2113 (0.0001)	0.4244 (0.0001)	0.1492 (0.0001)	0.1811 (0.0001)	-0.0721 (0.0086)	-0.0912 (0.0009)	-0.0030 (0.9144)	0.2047 (0.0001)	-0.0176 (0.5229)
Pq					0.2557 (0.0001)	0.0463 (0.0922)	0.0102 (0.7117)	-0.0038 (0.8909)	-0.0085 (0.7558)	0.0074 (0.7864)	-0.0317 (0.2480)	-0.0470 (0.0872)	0.0032 (0.9061)	0.0578 (0.0353)
Pm						-0.0574 (0.0365)	0.1327 (0.0001)	0.0984 (0.0003)	-0.0309 (0.2610)	-0.0779 (0.0045)	-0.0224 (0.4142)	-0.0472 (0.0858)	0.0998 (0.0003)	0.0321 (0.2421)
Pe							-0.3353 (0.0001)	-0.1411 (0.0001)	0.0376 (0.1715)	0.3118 (0.0001)	0.0189 (0.4921)	-0.1443 (0.0001)	-0.2399 (0.0001)	0.0511 (0.0628)
W1								0.4064 (0.0001)	-0.0495 (0.0717)	-0.2006 (0.0001)	-0.3106 (0.0001)	0.1170 (0.0001)	0.5860 (0.0001)	-0.0925 (0.0007)
W2									-0.0727 (0.0081)	-0.1108 (0.0001)	-0.1834 (0.0001)	0.0003 (0.9929)	0.3423 (0.0001)	-0.0084 (0.7592)
DIV										0.0120 (0.6617)	0.0232 (0.3982)	0.0034 (0.9016)	-0.0906 (0.0010)	0.0383 (0.1635)
D1											-0.2529 (0.0001)	-0.1351 (0.0001)	-0.1622 (0.0001)	-0.2371 (0.0001)
D2												-0.2457 (0.0001)	-0.2949 (0.0001)	-0.4311 (0.0001)
D3													-0.1576 (0.0001)	0.2304 (0.0001)
D4														-0.2765 (0.0001)

1. N = 1327 and PROB > |RHO| under HO:RHO = 0 appears in brackets; D5 = Ontario dummy.

The most interesting result is probably the significant negative correlation between the output volume per plant and output price (-0.13). This observation contradicts the hypothesis that, in an imperfect competition context, lower average costs of larger plants may reflect higher selling prices rather than lower costs in studies in which output is measured in value, not quantity.[27] Given that unit transportation costs for final products should be higher on average for larger plants, this negative relation between price net of transportation costs and plant size could be compatible with a quasi-unique market price, gross of transportation costs.[28] As indicated in Table 2 this negative correlation implied a higher negative coefficient of correlation between output volume and real unit cost (-0.16) than between output value and nominal unit cost (-0.10). This tends to indicate that assuming no output price differences among plants in a given industry will underestimate economies of scale.

Table 2 also indicates significant positive coefficient of correlations between real unit cost on one hand, and the materials and energy price indexes on the other. This result is expected from a price index correctly calculated in the sense that it supports the fact that input price

---

27. This criticism is quite common in the industrial economics literature (see Hazledine (1984), for example). However, advocates of this approach are not able to explain why small plants are not price-followers and do not set their price equal to the price set by larger plants. A related issue in this field is the interpretation of the typically-positive correlation obtained in empirical studies between market concentration and profitability, which suggests the existence of market power to some economists and efficiency to others. The negative correlation obtained between output volume and output price, combined with the large negative correlation between C/Q and C/V, provides some evidence supporting the efficiency interpretation since C/V could be interpreted as the inverse of the profit ratio including payments to capital services. (Note also the positive correlation between the implicit profit rate and output volume). However, the relatively smaller but statistically significant positive correlation between output price and profit tends to indicate that both factors are present, as found with U.S. data in recent studies (e.g. Martin (1988)). In any event, more formal tests should be developed and this is beyond the scope of this paper.

28. In the Manufacturing Census, the cost of materials and energy are inclusive of transportation costs, whereas the value of shipments includes own carriers' delivery expense for outward transportation but not common or contract carriers' charges.

differences reflect price differences for a given product rather than price differences for different products. In addition, combining this positive correlation between cost and input prices with the significant correlation between input prices and output volume suggests again that scale economy parameters would be biased when these input prices are not taken into account.

As indicated by previous studies (Miller (1981), Robidoux and Lester (1988)), these data indicate a strong positive relationship between plant size and wage rates. However, wage rates are not really correlated, on average, with production cost, as noted by Robidoux and Lester (1988). The same relation applies to product diversity; bigger plants produce, on average, more types of products, but more product diversity does not systematically affect costs.

Finally, prices seem to be significantly different, on average, from one region to another in some cases: output prices are higher in Ontario and lower in Prairies; materials prices are higher in B.C. and lower in the Prairies and the Atlantic provinces; energy prices are higher in Ontario and particularly in the Atlantic provinces but lower in B.C. and the Prairie provinces.[29]

VI. REGRESSION RESULTS

Based on the regression results, the CDR at 1/2 MES was calculated for the three alternative unit cost functions and the results are presented in Table 3.[30,31] In line with the correlation analysis, these results suggest

---

29. In general, the average correlations noted in this section hold in a majority of industries or, in some cases, in almost all industries.

30. Detailed regression results are presented in the annexes. Estimated output and regional dummy coefficients ($\hat{\alpha}_2$, $\hat{\alpha}_3$, and $\hat{\alpha}_7$ to $\hat{\alpha}_{10}$) not significantly different from zero at the 10 per cent confidence level (one-tail t-test) are not used in calculating CDR and all other statistics presented in this section.

that the CDR is substantially lower when the nominal form equation is used instead of the real or the full form equations. The main effect comes from the inclusion of deflated output, which increases the CDR by 70 per cent on average. In contrast, including input prices marginally reduces scale economies, but the combined effect of all prices increases the CDR by 68 per cent, on average.

TABLE 3

Cost Disadvantage Ratio (CDR) at 1/2 MES

		CDR at 1/2 MES			
	Industry	Nominal	Real	Full	Ratio[1]
1011	Slaughtering and Meat	0.00	3.34	2.23	-
1012	Poultry Processors	0.00	0.00	0.00	-
1020	Fish Products	1.47	0.00	0.00	0.00
1031	Fruits & Vegetables	4.43	1.28	5.13	1.16
1040	Dairy Products	0.00	4.16	4.07	-
1050	Flour & Cereal	0.00	24.99	22.86	-
1060	Feed	7.50	0.00	0.00	0.00
1071	Biscuits	0.35	15.14	15.16	43.31
1072	Bakeries	0.66	1.05	1.02	1.55
1089	Misc. Food Processors NES	7.44	6.00	4.94	0.66
1093	Breweries	0.80	1.42	1.12	1.40
1860	Carpet, Mat and Rug	10.92	5.25	5.87	0.54
2310	Hosiery	5.43	0.00	1.34	0.25
2392	Other Knitting	1.49	7.60	2.03	1.36
2513	Sawmills and Planing Mills	0.22	0.22	0.20	0.91
2520	Veneer and Plywood Mills	0.61	1.07	1.12	1.84
2710	Pulp and Paper Mills	8.25	9.41	10.66	1.29
2731	Corrugated Box	0.00	9.57	9.57	-
3550	Ready-Mix Concrete	1.87	2.06	1.21	0.65
3651	Petroleum Refining	0.40	2.75	2.23	5.58
	Weighted Average	2.65	4.50	4.49	1.68
	Unweighed Average	2.59	4.77	4.54	1.75

1. Ratio of full to nominal estimates.

31.(...continued)

31. In all tables of this section we present weighted and unweighed averages of industry results. The weighted average is calculated using the share of industry gross output in the sub-sample total gross output as weights.

On the other hand, a possibly undesirable but not surprising result is that the sectoral biases are highly variable. Although the CDR increases in a majority of industries, in 2 industries the CDR is virtually unaffected and in 6 others it decreases when full and nominal estimates are compared. Moreover, the variance of size of the bias is quite substantial in the 12 industries where the inclusion of prices raises the CDR.

An analysis of the CSA, which takes into account the distribution of plant size, does not change the main conclusion drawn from the analysis of CDR. In effect, as indicated in Table 4, including input and output prices increases the potential cost saving achievable by around 65 per cent at the national level. Not surprisingly, the results are similar to those for the CDR at the industry level, although some differences exist. In contrast to the CDR measure, however, the CSA contains some regional information suggesting that the bias in potential cost savings is higher in Quebec and Ontario and lower than the average in other regions when weighted averages are used (Table 6). The Atlantic provinces are the least affected by the inclusion of prices. However, the relative position of Ontario and Atlantic provinces are inverted when a simple average is used, which indicates that compositional effects are important in these two regions; the results are roughly stable in other regions.[32] Again industry-by-industry the results show a high degree of variability but, in all regions, 50 per cent of industries or more see their CSA increased when all prices are included and a decrease is observed in at most 33 per cent of industries.

The results also suggest that including input and output prices in the analysis increases the region-specific inefficiencies in all regions on average (Tables 5 and 6). This comes essentially from the inclusion of deflated output in the cost analysis which increases substantially the region-specific cost differences. In fact, including input prices does not change,

---

32. The unusual average plant size observed in our sample for Ontario and British Columbia should be also kept in mind in analyzing the relative position of regions, especially for these two provinces. (See Annex for more details).

TABLE 4

COST SAVING ACHIEVABLE BY ELIMINATING SCALE-RELATED INEFFICIENCY BY REGION - PERCENTAGE

	INDUSTRY	ATLANTIC			QUEBEC			ONTARIO			PRAIRIES			B.C.			CANADA		
		NOM.	REAL	FULL	NOM.	REAL	FULL	NOM.	REAL	FULL	NOM.	REAL	FULL	NOM.	REAL	FULL	NOM.	REAL	FULL
1011	MEAT PRODUCTS	X	X	X	0.0	5.9	3.8	0.0	0.0	0.0	0.0	2.7	1.1	0.0	3.0	5.8	0.00	2.42	1.45
1012	POULTRY PROCESSORS	X	X	X	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.00	0.00	0.00
1020	FISH PRODUCTS	10.6	0.0	0.0	9.0	0.0	0.0	X	X	X	-	-	-	X	X	X	10.45	0.00	0.00
1031	FRUIT & VEGETABLES	X	X	X	12.0	9.9	14.0	1.1	0.0	1.7	X	X	X	9.5	8.8	13.9	4.75	2.93	5.63
1040	DAIRY PRODUCTS	0.0	16.1	20.3	0.0	2.9	2.5	0.0	6.9	6.6	0.0	5.6	5.2	X	X	X	0.00	5.49	5.34
1050	FLOUR & CEREAL	X	X	X	0.0	29.1	32.3	0.0	2.1	8.2	0.0	19.2	19.5	-	-	-	0.00	10.85	14.74
1060	FEED	7.0	0.0	0.0	6.1	0.0	0.0	1.3	0.0	0.0	10.2	0.0	0.0	1.2	0.0	0.0	4.65	0.00	0.00
1071	BISCUITS	X	X	X	0.0	18.3	23.7	0.0	7.5	11.5	X	X	X	-	-	-	0.66	11.36	16.72
1072	BAKERIES	9.8	0.0	0.0	3.7	0.0	0.0	6.5	7.2	6.4	5.6	3.8	2.6	X	X	X	5.53	2.28	1.81
1089	MISC. FOOD	X	X	X	0.0	7.7	3.7	11.9	9.9	7.3	10.8	10.0	7.7	6.4	12.7	9.0	6.83	9.68	6.49
1093	BREWERIES	8.3	13.8	9.3	0.0	0.0	0.0	0.0	0.0	0.0	3.6	0.0	5.2	0.9	4.1	2.0	1.03	1.07	1.44
1860	CARPET, MAT & RUG	-	-	-	15.4	22.5	24.5	7.5	14.9	16.4	X	X	X	-	-	-	11.82	19.22	21.04
2310	HOSIERY	-	-	-	2.3	0.0	0.0	4.5	0.0	5.0	-	-	-	-	-	-	3.18	0.00	1.66
2392	OTHER KNITTING	X	X	X	5.2	24.8	5.2	X	X	X	0.8	10.6	2.1	-	-	-	4.57	17.90	4.71
2513	SAWMILLS & PLANING	7.2	11.0	16.2	5.2	1.1	0.0	2.0	3.7	3.4	3.1	0.0	0.0	3.3	2.3	2.2	3.64	2.42	2.35
2520	VENEER & PLYWOOD	X	X	X	10.3	10.5	10.2	5.5	12.3	11.0	0.0	12.7	3.2	0.7	11.3	7.9	2.27	11.54	8.29
2710	PULP AND PAPER	3.6	11.1	7.8	2.6	8.1	6.0	7.2	7.8	9.0	0.0	0.0	4.3	5.2	4.4	5.2	4.34	7.13	6.64
2732	CORRUGATED BOX	0.0	9.0	9.2	0.0	7.2	7.2	0.0	5.4	5.4	0.0	11.0	10.9	0.0	0.0	0.0	0.00	5.86	5.84
3550	READY-MIX CONCRETE	8.3	16.9	8.6	6.8	9.1	5.0	4.4	5.7	1.9	0.0	2.7	0.0	2.7	19.8	5.1	3.67	6.58	2.46
3651	PETROLEUM REFINING	0.0	0.0	0.0	0.2	2.6	1.6	0.0	1.9	1.9	1.7	0.0	0.6	0.0	4.6	5.2	0.38	1.64	1.55
	WEIGHTED AVERAGE	2.9	4.4	3.7	2.0	5.4	4.0	2.2	3.9	4.1	1.4	2.3	2.2	3.1	4.3	4.3	2.30	4.20	3.80
	UNWEIGHTED AVERAGE	5.6	9.5	10.3	3.9	8.0	7.0	3.8	5.2	5.3	4.1	6.1	6.2	2.8	5.1	4.1	3.39	5.92	5.41

X - CONFIDENTIAL; - - NO PLANTS IN THE SAMPLE

TABLE 5

COST SAVING ACHIEVABLE BY ELIMINATING REGION-RELATED INEFFICIENCY BY REGION - PERCENTAGE

	INDUSTRY	ATLANTIC			QUEBEC			ONTARIO			PRAIRIES			B.C.			CANADA		
		NOM.	REAL	FULL	NOM.	REAL	FULL	NOM.	REAL	FULL	NOM.	REAL	FULL	NOM.	REAL	FULL	NOM.	REAL	FULL
1011	MEAT PRODUCTS	X	X	X	0.0	0.0	0.0	0.0	19.0	14.6	3.0	7.9	7.6	0.0	18.8	13.8	1.07	10.74	8.75
1012	POULTRY PROCESSORS	X	X	X	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.21	0.26	0.73
1020	FISH PRODUCTS	0.0	24.9	26.0	0.0	30.6	34.4	X	X	X	-	-	-	X	X	X	0.00	23.89	25.11
1031	FRUIT & VEGETABLES	X	X	X	0.0	0.0	0.0	0.0	6.8	7.4	X	X	X	9.7	17.1	18.7	1.51	7.29	7.94
1040	DAIRY PRODUCTS	0.0	6.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	X	X	X	0.00	0.29	0.00
1050	FLOUR & CEREAL	X	X	X	14.6	8.2	0.0	17.8	11.1	0.0	15.1	0.0	12.2	-	-	-	16.24	7.49	3.54
1060	FEED	0.0	13.5	0.0	0.0	7.4	0.0	0.0	7.5	0.0	0.0	0.0	5.4	9.3	7.1	0.0	1.01	6.76	0.67
1071	BISCUITS	X	X	X	0.0	0.0	0.0	0.0	0.0	0.0	X	X	X	-	-	-	0.00	1.90	0.00
1072	BAKERIES	15.4	19.8	25.6	16.7	12.6	13.0	16.6	19.3	24.7	9.2	2.8	2.1	X	X	X	12.05	10.53	12.29
1089	MISC. FOOD	X	X	X	17.6	15.9	10.0	14.8	15.8	9.7	0.0	0.0	0.0	5.6	6.8	2.3	14.46	13.25	8.02
1093	BREWERIES	13.4	18.3	14.1	28.0	20.5	21.9	0.0	0.0	0.0	15.2	22.9	0.0	2.6	19.3	19.9	15.22	13.81	10.69
1860	CARPET, MAT & RUG	-	-	-	0.0	0.0	0.0	0.0	0.0	0.0	X	X	X	-	-	-	0.00	0.85	0.83
2310	HOSIERY	-	-	-	0.0	0.0	0.0	0.0	21.3	25.1	-	-	-	-	-	-	0.00	7.06	8.34
2392	OTHER KNITTING	X	X	X	0.0	0.0	16.9	X	X	X	0.0	22.8	22.6	-	-	-	0.00	4.95	11.35
2513	SAWMILLS & PLANING	5.4	15.3	14.0	0.0	13.2	16.7	0.0	16.3	15.9	0.0	0.0	17.7	0.0	1.1	0.0	0.21	4.97	5.44
2520	VENEER & PLYWOOD	X	X	X	0.0	17.4	18.6	0.0	0.0	0.0	0.0	0.0	12.8	1.6	17.0	18.5	8.18	13.67	15.75
2710	PULP AND PAPER	18.6	17.5	21.0	15.5	13.9	13.5	14.6	18.0	12.9	16.9	20.4	14.0	0.0	0.0	0.0	11.71	11.76	10.64
2732	CORRUGATED BOX	0.0	10.5	11.3	3.0	16.8	17.2	0.0	10.3	11.1	0.0	0.0	0.0	0.0	11.3	12.1	0.57	10.65	11.26
3550	READY-MIX CONCRETE	9.4	21.1	16.5	0.0	7.3	6.5	0.0	0.0	0.0	0.0	21.5	8.6	0.0	0.0	0.0	0.40	8.40	4.50
3651	PETROLEUM REFINING	6.5	9.9	9.5	0.0	6.6	4.9	6.5	6.6	4.9	2.7	0.0	0.0	1.7	3.0	1.8	3.51	5.48	4.37
	WEIGHTED AVERAGE	8.7	14.2	14.3	6.1	8.4	7.8	5.4	10.4	7.7	3.9	5.1	4.6	2.0	3.2	2.5	5.09	8.10	6.96
	UNWEIGHTED AVERAGE	5.9	16.6	12.3	4.8	8.5	8.7	3.5	9.5	8.0	3.4	8.9	8.2	4.7	6.8	5.8	4.32	8.20	7.51

X - CONFIDENTIAL; - - NO PLANTS IN THE SAMPLE

or more often reduced, as expected, the region-related inefficiencies. Using weighted averages, the potential region-related cost savings increase by 37 per in Canada moving from nominal to full estimates. There is not a very high degree of regional variability, varying from 17 per cent in Prairies to 65 per cent in Atlantic provinces (Table 6). However, using the unweighed average increases these biases to 100 per cent and more in all regions except British Columbia, indicating important composition effects. In other words, in our sample, industries with highly underestimated region-specific inefficiencies are relatively small in each of these regions. Finally, as for scale results, the sectoral results are characterized by a high degree of variability. The only notable exception is the Atlantic provinces where the region-specific inefficiencies are stable or increase in all industries analyzed.

TABLE 6

Ratio of Full to Nominal Aggregate Estimates[1]

	Atlantic	Quebec	Ontario	Prairies	B.C.	Canada
Scale-Related CSA						
Weighted Average	1.26	1.98	1.86	1.50	1.39	1.65
Unweighed Average	1.86	1.77	1.39	1.54	1.45	1.60
Region-Related CSA						
Weighted Average	1.65	1.29	1.41	1.17	1.25	1.37
Unweighed Average	2.07	1.82	2.28	2.38	1.24	1.74

1. Source: Tables 4 and 5.

If the input price indexes are correctly measured, in the sense that they capture variation in price for a homogeneous composite input, then estimated coefficients of these prices are expected to be positive. This is effectively the case for the materials price coefficient, which is positive in almost all industry regressions and statistically significant in two-thirds of cases as indicated in Table A6 of the Annex. The results are less convincing

for the energy price coefficient, which is generally not statistically significant and sometimes negative. The poor quality of data used to calculate this energy price index is the most important factor in explaining these results. (See the Annex for more details.)

Similarly, if input and output prices are correctly measured, then including these prices in the regression analysis should have positive impact on statistical results such as the goodness-of-fit, the shape of the cost curve and the significance of returns to scale. As a first indication, Annex Table A1 indicates that the adjusted R^2 increases in most of the 20 industries when prices are included and, as a result, the average adjusted R^2 increases from 0.08 to 0.26 and to 0.33 when output price and all prices are taking into account, respectively. Moreover, the nominal equation produces unexpected or bizarre cost curve shapes in three industries that are eliminated when prices are included.[33] Since only output coefficients significantly different from zero at a 10 per cent confidence level[34] or better are kept in any calculation, it is hard to compare the statistical significance of returns to scale for the three functional forms. However, results presented in Tables A4 to A6 indicate that more output coefficients are significant at a 1 per cent confidence level when prices are included, particularly output price, than otherwise.

The results obtained for wage rate coefficients with the nominal form are similar to results obtained by Robidoux and Lester (1988): most of them are not statistically significant. However, coefficients on W1 and W2 are more often positive (75 per cent), particularly for W1, whereas it was the inverse in Robidoux and Lester. When prices are taken into account, the general figure is similar with the important exception that W1 has a significant positive impact on costs in half of the industries (see Tables A4 to A6). The number of significant regional dummy coefficients increases substantially when price indexes are included, implying higher differences in

---

33. There are U-shaped, inverted U-shaped and inverted hyperbolic curves in industries 3550, 1050 and 1012, respectively (see Annex Table A1).

34. One tail t-test.

region-specific cost efficiency as noted above. On the other hand, the coefficient of the product diversity index is not systematically affected by the price indexes.

Up to now, we have compared results obtained with the three functions and noted differences without formally testing if these estimates are statistically different. The tests of the difference between the nominal and the full estimates, based on the approach proposed by Hausman (1978), are presented in Annex Table A7. The results indicate that these differences are often significant. More importantly, coefficient differences implicitly used in calculating the CDR and CSA measures are almost all significant.[35] In other words, based on the Hausman test, the impact noted in this section of input and output prices on scale economies and regional cost efficiencies are statistically significant.[36]

VII. CONCLUSION

It is quite common in multi-industry cross-section studies estimating cost or production functions to assume that input and output prices are the same for all plants in a given industry. This study tests this hypothesis and its impact on cost function estimates with Canadian data.

The results obtained suggest that input and output prices vary among plants in a given four-digit industry. Moreover, these prices are correlated with basic industrial variables such as output, unit cost and region. We have found that, in general, larger plants sell their products,

---

35. For the scale parameters, the exceptions are industries 1031, 2392 and 2513, where only marginal changes are observed. Similarly, the few changes in regional dummy coefficients implicitly used in calculating the regional cost efficiencies that are not statistically significant are generally small.

36. Given the large data manipulation of various variances-covariances matrices, no test of the overall change in coefficients was done. However, these global tests would not likely change the basic conclusion.

measured in homogeneous units, at lower prices than smaller plants. This result is in contradiction with the common idea in industrial economics that large plants may obtain a higher price for their products in imperfectly competitive markets. On the other hand, our empirical finding could be the logical result of the standard analysis of transportation costs which predicts that unit transport costs increase with distance and then, with plant size, ceteris paribus. In other words, the negative correlation between output volume and output price may result in a quasi-unique price gross of transportation costs. More analysis is necessary to investigate this issue.

As a corollary, the results also suggest that explicitly taking into account the output price variance substantially increases estimates of economies of scale and region-specific inefficiencies. In contrast, intermediate input price variance marginally reduces scale economies and region-specific inefficiencies, but the impact of output price generally dominates. Even if this average picture holds in most of industries, some industry results are not affected or are affected in the other way when we relax the constant price hypothesis. However, these disparities in industry results are not surprising considering the potential biases involved.

References

Allen, R.G.D. (1975) Index Numbers in Theory and Practice, The MacMillan Press Ltd., London

Baldwin, J.R., and Gorecki, P.K., (1986), The Role of Scale of Canada/U.S. Productivity Differences in the Canadian Manufacturing Sector, 1970-1979; Royal Commission on the Economic Union on Development Prospects for Canada, University of Toronto Press.

Besley, D.A., Kuh, E. and R.E. Welsh (1980) Regression Diagnotics: Identifying Influential Data and Sources of Collinearity, John Wiley & Sons.

Caves, D.W. Christensen, L.R. and W.E. Diewert, (1982), "Multilateral comparisons of output, input, and productivity using superlative index numbers", The Economic Journal, pp. 73-86.

Corbo, V. and Meller, P., (1979), "The Translog Production Function: Some Evidence From Establishment Data", Journal of Econometrics, Vol. 10, No. 2, pp. 193-200.

De Brabander and Vanlommel, (1978), "Economies of Scale, Minimum Optimal Plant Size and Effectiveness of Market Structures in Belgian Industry: Anno 1970", European Economic Review, 11, pp. 363-377.

Diewert, W.E. (1976), "Exact and superlative index numbers", Journal of Econometrics, vol. 4, pp. 115-145.

Fisher, I. (1922), The Making of Index Numbers, (Boston: Houghton Miftin Company, 1922), third edition (1927) was reprinted in 1967 by Augustus M. Kelly, Publishers, New York.

Fuss, M. (1977), "The demand for energy in Canadian manufacturing: An example of the estimation of production structures with many inputs", Journal of Econometrics, 89-116.

Fuss, M.A., et Gupta, V.K., (1981), "A Cost Function Approach to the Estimation of Minimum Efficient Scale, Returns to Scale, and Sub-Optimal Capacity", European Economic Review, 15, pp. 123-135.

Griliches, Z. et Ringstad, V., (1971), Economies of Scale and the Form of the Production Function, Contribution to Economic Analysis No. 72, North-Holland, Amsterdam.

Gupta, V.K., (1968), "Cost Functions, Concentration, and Barriers to Entry in Twenty-nine Manufacturing Industries of India", The Journal of Industrial Economics, Vol. XVII, No. I, pp. 57-72.

Hausman, J.A. (1978), «Specification Test in Econometrics», Econometrica, 46, 6, pp. 1251-1291.

Hazledine, T., (1984), Trade Industrial Policy and Canadian Manufacturing: The Harris/Cox Model and an Alternative, University of British Columbia, mimeo.

Hulten, C.R. (1973), "Divisia index numbers", Econometrica, Vol. 41, No. 6.

Maddala, G.S. (1977), Econometrics, McGraw-Hill.

Martin, S. (1988), "Market Power and/or Efficiency?", The Review of Economics and Statistics, pp. 331-335.

Miller, E.M. (1981), "Variation of Wages Rates with Size of Establishment", Economics Letters, pp. 281-286.

Robidoux, B., (1987), Économies d'échelle: Un survol des techniques d'estimation et des problèmes méthodologiques, Department of Finance Working Paper #87-1, Canada.

Robidoux, B. and J. Lester, (1988), Econometric Estimates of Scale Economies in Canadian Manufacturing, Department of Finance Working Paper #88-3, Canada.

Theil, H. (1973), "A new index number formula". The Review of Economics and Statistics, 498-502.

Törnqvist, M.A. (1936), "The Bank of Findland's Consumption Price Index", in Collected Scientific Papers of Leo Törnqvist, the Research Institute of the Finnish Economy, Helsinki, 1981.

White, H. (1980), «A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity», Econometrica, 48, 4, pp. 817-838.

ANNEX

A.1 CONSTRUCTION OF PRICE INDEXES AND SELECTION OF THE FINAL SAMPLE

The typical plant in a given manufacturing industry produces a variety of products and uses different types of materials and energy inputs. Data on quantities and values for these inputs and outputs are available in the basic Census files.[37] As noted in the text, these data could be used to calculate plant-specific Divisia-Wald price indexes.

However, some analysis of these data must be done before they can be used for this exercise. First, we have verified that each quantity of a given output or input was based on the same unit of measure. Specific quantities based on an unusual unit of measure were simply removed from the sample.[38] Second, for each plant and variable, the sum of detailed values were compared to the total value. With this procedure, a limited number of keypunch errors were detected and corrected. Third, specific outputs or inputs without quantity or value were excluded. After these missing data points were removed, we calculated a plant-specific price for each output and input existing in a given industry (i.e., the p_i^j of equation 14). Examination of theses prices (and their related dispersion statistics) in conjunction with questionnaires permitted the identification of outliers and residual categories that were too large, usually labelled "other products....". Based on this analysis, a substantial number of plant data points and industry-specific input or output were removed from the sample. Finally, in order to obtain representative price indexes, plants that had less than 70 per cent of their original input *OR* output values after this selection process were excluded from the final calculation.

Using this sample, Divisia-Wald price indexes were calculated for output, materials and energy and for each plant. This price data base was then merged with another data base including all other relevant variables. This last sample corresponds also to a sub-sample of the Census universe since some plants are excluded. These plant exclusions reflect concerns about measurement error and the quality of the data reported by small plants. Detailed exclusion criteria can be found in Robidoux and Lester (1988). The resulting sample contains 38 four-digit industries with all variables available for at least 15 plants which was the lower bound we fixed for the regression analysis given degree-of-freedom requirements.

The Table A1 presents selected pre- and post-regression results obtained for these industries. These preliminary results have been used in the selection of the final sample. Many problems were considered, the most important of which is the relative homogeneity of basic products used to

---

37. Each industry questionnaire contains the same ten types of energy inputs. However, the output and materials types are industry-specific and they vary in number from one questionnaire to another. For most industries, the number of input and output types is between 50 and 100.

38. Only a marginal number of observations was removed by this procedure.

TABLE A1

SELECTION OF THE FINAL REGRESSION SAMPLE : PRELIMINARY REGRESSION RESULTS AND DATA INFORMATION

	INDUSTRY	RETURNS TO SCALE(1)			ADJUSTED R^2			CONDITION NUMBER(2)			ESTIMATED COEFFICIENTS OF(3)	
		NOMINAL	REAL	FULL	NOMINAL	REAL	FULL	NOMINAL	REAL	FULL	ENERGY PRICE	MATERIALS PRICE
1011	SLAUGHTERING AND MEAT PROCESSORS	C	I	I	0.01	0.33	0.49	24 (W)	24 (W)	42 (E)	0.19 (5)	0.51 (1)
1012	POULTRY PROCESSORS	D	C	C	0.30	0.19	0.21	35 (W)	35 (W)	96 (EM)	-0.08	0.37 (10)
1020	FISH PRODUCTS	I	C	C	0.03	0.24	0.24	38 (WO)	37 (WO)	52 (E)	-0.02	0.14 (5)
1031	FRUIT & VEG. CANNERS & PRESERVERS	I	I	I	0.37	0.26	0.29	20 (W)	20 (W)	37 (E)	0.26 (5)	0.12
1032	FROZEN FRUIT & VEGETABLE PROCESSORS	IU	C	C	-0.55	-0.61	-1.17	27 (HOW)	27 (HOW)	114 (ALL)	0.84	-1.29
1040	DAIRY PRODUCTS	C	I	I	0.07	0.08	0.06	20 (W)	20 (W)	48 (E)	0.09	0.00
1050	FLOUR & BREAKFAST CEREAL PRODS.	IU	I	I	0.45	0.07	0.38	64 (W)	62 (W)	74 (WME)	0.03	0.89 (1)
1060	FEED (4)	I	C	C	0.21	0.13	0.56	30 (W)	31 (W)	61 (M)	-	0.98 (1)
1071	BISCUITS	I	I	I	-0.36	0.30	0.15	63 (ALL)	56 (ALL)	267 (ALL)	-0.83	0.43
1072	BAKERIES	I	I	I	0.07	0.40	0.43	26 (W)	25 (W)	43 (EM)	-0.35	0.49 (1)
1081	CONFECTIONERY	I	D	D	0.22	0.13	0.06	32 (ALL)	34 (ALL)	78 (MED)	0.46	0.57 (5)
1089	MISC. FOOD PROCESSORS NES	I	I	I	0.02	0.21	0.29	18 (W)	18 (W)	37 (ME)	0.36 (1)	0.17
1092	DISTILLERIES	D	U	D	0.55	0.56	0.56	45 (W)	47 (W)	59 (WE)	0.77 (5)	-0.16
1093	BREWERIES (5)	I	I	I	0.50	0.70	0.71	52 (W)	50 (W)	70 (WM)	0.07	0.32 (10)
1094	WINERIES	I	C	C	-0.03	0.26	-0.21	110 (ALL)	103 (ALL)	182 (ALL)	-0.13	-0.14
1620	WOOL YARN AND CLOTH	I	C	C	0.15	0.15	0.17	35 (W)	34 (W)	89 (EW)	0.41	-0.58 (10)
1860	CARPET, MAT AND RUG	I	I	I	-0.13	0.59	0.47	59 (ALL)	59 (WO)	107 (ALL)	-0.13	-0.11
2310	HOSIERY (4),(5)	I	C	I	-0.01	0.20	0.48	23 (W)	23 (W)	29 (W)	-	0.48 (1)
2392	OTHER KNITTING	I	I	I	0.02	0.04	0.00	35 (W)	32 (W)	69 (E)	0.31	0.15
2431	MEN'S CLOTHING	C	C	C	0.06	0.02	0.07	20 (W)	20 (W)	41 (E)	0.60 (1)	0.08
2441	WOMEN'S CLOTHING	C	I	I	0.02	0.25	0.29	19 (ALL)	17 (ALL)	41 (E)	-0.59 (1)	0.05
2450	CHILDREN'S CLOTHING	C	I	C	-0.13	0.05	0.04	20 (W)	20 (W)	44 (E)	-0.08	0.15 (10)
2511	SHINGLE MILLS	C	C	C	0.12	0.28	0.18	55 (OO)	81 (OO)	113 (ALL)	0.25	-0.01
2513	SAWMILLS AND PLANING MILLS	I	I	I	0.02	0.05	0.24	25 (W)	25 (W)	39 (E)	0.07	0.31 (1)
2520	VENEER AND PLYWOOD MILLS (4)	I	I	I	0.02	0.22	0.29	28 (W)	27 (W)	31 (W)	-	0.31 (1)
2710	PULP AND PAPER MILLS	I	I	I	0.30	0.31	0.32	31 (W)	31 (W)	48 (ME)	-0.01	0.20 (5)
2732	CORRUGATED BOX	C	I	I	-0.04	0.22	0.19	31 (W)	30 (W)	75 (M)	-0.05	0.09
2740	MISCELLANEOUS PAPER CONVERTERS	C	U	U	0.02	0.06	0.08	22 (W)	22 (W)	50 (E)	0.38	0.31 (10)
2910	IRON AND STEEL MILLS	I	C	IU	-0.17	-0.27	-0.38	35 (W)	35 (W)	64 (W)	-0.01	-0.30
2960	ALUMINUM ROLLING, CASTING ETC.	U	U	U	0.75	0.40	0.40	28 (H)	28 (H)	97 (MOE)	0.45 (1)	-0.09
2980	METAL ROLLING, CASTING ETC. NES	D	C	D	0.31	-0.04	0.09	27 (W)	27 (W)	48 (ME)	0.91 (1)	0.09
3020	FAB. STRUCTURAL METAL	C	C	C	-0.34	-0.05	-0.10	36 (ALL)	34 (ALL)	63 (EW)	0.86 (5)	0.18
3050	WIRES AND WIRES PRODUCTS	I	I	I	-0.31	-0.37	0.07	28 (H)	28 (W)	56 (E)	-1.12 (1)	0.75 (1)
3541	CONCRETE PIPE	I	I	C	0.55	0.32	0.16	66 (DW)	67 (DW)	108 (ALL)	0.75	-1.03
3550	READY-MIX CONCRETE	U	I	I	0.22	0.41	0.50	29 (W)	28 (W)	39 (MOW)	0.10	0.59 (1)
3651	PETROLEUM REFINING (4),(5)	I	I	I	-0.07	0.16	0.29	34 (W)	35 (W)	126 (WM)	-	0.71 (5)
3730	PLASTICS & SYNTHETIC RESINS	IU	IU	IU	0.26	0.39	0.23	43 (H)	43 (H)	189 (ALL)	1.16	-0.71
3782	INDUSTRIAL CHEM. (INORGANIC) NES	IU	IU	IU	0.21	0.12	0.35	29 (W)	29 (W)	43 (W)	0.74 (1)	-0.28 (10)

(1) I - increasing returns to scale for all output levels; C - constant returns to scale; D - decreasing returns to scale for all output levels; U - U-shaped unit cost curve; IU - inverted U-shaped unit cost curve. Only output coefficients significantly different from zero at the 10% confidence level for a one-tail t-test are used.

(2) Mnemonics between brakets indicate variables potentially affected by collinearity with other exogenous variables; i.e., estimated coefficients which have a large share of their variances associated with condition numbers higher than 30 or the highest one. W - wage rates; H - product-diversity index; O - output variables; D - regional dummies; E - energy price; M - materials price; ALL - all variables. The intercept estimate is affected in almost all regressions.

(3) (i) indicates significantly different from zero at the i % confidence level for a one-tail t-test.

(4) The energy price is not included in these regressions since pre-regessions suggest statistically significant negative coefficients.

(5) The product diversity index was removed from these regressions since it was too highly correlated with other exogenous variables.

TABLE A1 (CONT'D)

SELECTION OF THE FINAL REGRESSION SAMPLE : PRELIMINARY REGRESSION RESULTS AND DATA INFORMATION

	INDUSTRY	HOMOGENEITY OF PRODUCT-LINE		PER CENT OF ESTIMATED QUANTITY OR VALUE(2)			PER CENT KEPT OF		LIST OF MAJOR PROBLEMS(3)	INDUSTRY SELECTED
		OUTPUT	MATERIALS	ENERGY	MATERIALS	OUTPUT	OUTPUT RANGE	INDUSTRY OUTPUT		
1011	SLAUGHTERING AND MEAT PROCESSORS	M	M	26.0	7.0	5.7	99.8	65.1	-	x
1012	POULTRY PROCESSORS	M	G	45.0	5.4	6.6	100.0	71.6	-	x
1020	FISH PRODUCTS	M	M	40.0	4.9	4.0	10.2	25.6	D	x
1031	FRUIT & VEG. CANNERS & PRESERVERS	M	M	20.1	4.6	4.4	98.7	42.4	-	x
1032	FROZEN FRUIT & VEGETABLE PROCESSORS	M	M	22.6	23.2	37.7	-	-	E,F,M,C	
1040	DAIRY PRODUCTS	M	G	35.9	5.7	8.8	54.6	45.7	-	x
1050	FLOUR & BREAKFAST CEREAL PRODS.	M	M	13.2	2.4	3.1	65.2	54.1	-	x
1060	FEED	M	M	56.4	9.6	7.7	66.2	36.9	-	x
1071	BISCUITS	M	M	17.0	6.1	1.8	57.8	42.5	C	x
1072	BAKERIES	M	M	45.8	17.4	11.1	43.3	20.8	E,D	x
1081	CONFECTIONERY	M	M	19.8	4.4	0.7	-	-	S,F,C	
1089	MISC. FOOD PROCESSORS NES	M	M	23.4	2.2	2.3	99.8	40.5	-	x
1092	DISTILLERIES	M	M	10.8	1.6	2.3	-	-	S	
1093	BREWERIES	G	M	12.6	0.5	0.0	98.2	83.8	-	x
1094	WINERIES	M	B	9.7	1.3	2.6	-	-	S,F,C	
1820	WOOL YARN AND CLOTH	M	M	20.6	7.8	10.0	-	-	M	
1860	CARPET, MAT AND RUG	G	M	11.4	3.3	3.7	94.5	87.3	C	x
2310	HOSIERY	M	M	30.5	17.4	15.4	89.8	54.1	E	x
2392	OTHER KNITTING	M	M	42.9	14.9	12.2	100.0	47.2	E	x
2431	MEN'S CLOTHING	B	M	73.7	29.4	19.8	-	-	H	
2441	WOMEN'S CLOTHING	B	M	89.0	29.5	18.2	-	-	H	
2450	CHILDREN'S CLOTHING	B	M	89.9	36.3	22.9	-	-	H	
2511	SHINGLE MILLS	M	B	74.9	25.0	15.3	-	-	H,C,S,E	
2513	SAWMILLS AND PLANING MILLS	G	M	37.5	2.0	2.2	100.0	64.6	-	x
2520	VENEER AND PLYWOOD MILLS	G	G	26.6	6.6	6.4	94.0	78.1	-	x
2710	PULP AND PAPER MILLS	G	M	3.0	0.6	0.3	99.5	85.5	-	x
2732	CORRUGATED BOX	G	G	22.8	16.6	8.4	84.8	87.8	-	x
2740	MISCELLANEOUS PAPER CONVERTERS	M	M	29.9	27.3	23.2	-	-	S,F,E	
2910	IRON AND STEEL MILLS	G	M	2.0	6.0	2.3	-	-	F,S	
2960	ALUMINUM ROLLING, CASTING ETC.	G	M	11.4	5.6	4.7	-	-	S,F	
2980	METAL ROLLING, CASTING ETC. NES	G	M	47.7	31.4	33.3	-	-	S,F,E	
3020	FAB. STRUCTURAL METAL	M	M	67.7	38.4	40.0	-	-	C,S,F,E	
3050	WIRES AND WIRE PRODUCTS	M	M	26.8	14.7	16.3	-	-	N,E,F	
3541	CONCRETE PIPE	G	M	58.0	26.4	29.0	-	-	C,S,F,E	
3550	READY-MIX CONCRETE	M	M	38.5	15.9	14.0	100.0	49.5	E	x
3651	PETROLEUM REFINING	M	M	1.2	0.0	0.0	96.6	82.1	-	x
3730	PLASTICS & SYNTHETIC RESINS	G	M	5.7	1.8	1.7	-	-	S,C,M	
3782	INDUSTRIAL CHEM. (INORGANIC) NES	G	M	8.0	4.2	1.0	-	-	M,S	
	SUB-TOTAL			11.3	5.2	4.2				
	TOTAL MANUFACTURING			14.7	9.0	5.2				

(1) B, M and G indicate bad, medium and good homogeneity of product-lines, respectively. These results are based on ICC codes analysis and questionnaire examinations.

(2) Percentage of energy, materials and output value respectively, which has quantity or value estimated by Statistics Canada in the base universe sample.

(3) S - cost curve shapes; F - goodness-of-fit; C - collinearity diagnostics, H - homogeneity of product-lines; E - per cent of estimated data used in calculating output, materials and energy prices; M - materials price coefficient; N - energy price coefficient; D - per cent of data kept in the regression sample.

calculate price indexes. Homogeneity is a sine qua non condition to obtain a correctly measured price index. The examination of basic price variance, commodity codes and questionnaires permit an evaluation of the average product-homogeneity in a given industry. In general terms, products seem to be more homogeneous in sectors producing industrial goods and less in sectors producing consumption goods. Heterogeneity of output categories were crucial in rejecting clothing industries (2392, 2431, 2441) as indicated in Table A1.

Another important problem comes from the fact that Census officers estimate some value or quantity for product-lines not reported by manufacturers. Although industry average prices seem to be mostly used as estimates, this procedure could create obvious problems related to proxy measurement, especially when a large share of plant data or of industry data are estimated. Column 6 of Table A1 shows that, on average, this problem is more important for energy data and less for materials and output data, respectively. In 11 industries not selected in the final sample this problem was particularly present.[39] In some industries, we also obtained negative coefficients for energy or/and materials prices. In four industries (1060, 2310, 2520 and 3651) where the only important problems were data quality and statistically significant negative coefficients for the energy price, this variable was simply excluded from regressors. In contrast, when this problem of poor quality data and coefficient signs for materials or energy was present in conjunction with other problems, the industry was removed from the sample.

The highest condition number, which is an indicator of collinearity between independent variables, was also an important factor of rejection. In addition to the highest condition number, the estimated coefficients that have a large share of their variances associated with condition numbers higher than 30 or the highest one are indicated in brackets (see panel 3 of Table A1). As we know from the seminal work of Besley et al. (1980), these coefficients and their variances could be seriously affected by multicollinearity. Industry results that seemed to be excessively affected by this problem were rejected with special attention given to output coefficients.

The goodness-of-fit, as indicated by the adjusted R^2, was also used as selection criteria for the final sample. Industry results are not really interesting and indeed could be misleading when the fit of data to the function is very poor. Consequently, regression results with negative or very low adjusted R^2 were rejected. Similarly, industries having bizarre unit cost curve shapes such as inverted hyperbolic, inverted-U or even U-shaped were identified as potential problems. In general, when this problem was present in conjunction with other problems, such as multicollinearity and poor fit, the industry was excluded from the sample.

---

39. Given their relative importance in the regression work, more weight in the rejection decision was given to output data and less to input data. The regression results for the final sample should then be interpreted in light of this procedure, especially for the energy price where the quality of the underlying data is poor.

A summary list of problems and the industries selected is presented in the last two columns of Table A1. As we can see, in most industries rejected, more than one problem was present. After taking into account all these problems, 18 industries were excluded, leaving a final sample of 20 industries.

A.2 REPRESENTATIVENESS OF THE FINAL SAMPLE

Is the sub-sample of manufacturing industries used in this analysis representative of overall manufacturing? This annex provides some background information on this important question. As indicated in Table A2, the 20 industries selected in this study correspond to roughly one third of manufacturing production in 1979. On a regional basis, this proportion is higher in all regions but Ontario, varying from 35 per cent in Quebec to roughly 70 per cent in British Columbia and the Atlantic provinces. In contrast, only 20 per cent of Ontario manufacturing output is covered by the sample. This reflects the fact that some important Ontario manufacturing industries, such as motor vehicles, motor vehicle parts and iron and steel mills, are not in the sample. However, most major 4-digit Canadian or regional manufacturing industries are present in the sample: petroleum refining; pulp and paper mills; sawmills and planing mills; slaughtering and meat processors; fish products; dairy products.

Another important sample characteristic is the data exclusions applied to the basic data base. As shown in Table A2, around one third of the gross output of the 20 industries in the sample is removed from the final sample as a result of all the exclusion criteria applied in this study. This is not trivial. In general, however, industry-by-industry the gross output coverage is in general over 40 per cent (Table A1).[40] More importantly, the output range -- i.e. maximum minus minimum output -- is not affected too seriously in most industries. (The fish products industry is a notable exception.)

Table A3 gives details on the regional dispersion of unit cost and plant size in the sample compared to the whole Census sample. The regional dispersion of unit cost in the sample is roughly the same as for the whole manufacturing, although some significant differences seem to exist. In terms of average gross output per plant, some significant differences are obvious in the Ontario and British Columbia. Plants in Ontario tends to be smaller in the final sample than for the Census and vice versa for British Columbia. As a result, it would be reasonable to interpret the regional scale results with caution, particularly for these two regions.

---

40. Fish products and biscuits are notable exceptions.

TABLE A2

Gross Output Share by Region in the Regression Sample[1]

	Industry	Atlantic	Quebec	Ontario	Prairies	B.C.	Canada
				- percentage - [2]			
1011	Slaughtering and Meat Processing	x	7.9	17.5	32.4	2.2	12.3
1012	Poultry Processors	x	2.7	2.8	1.0	1.3	2.0
1020	Fish Products	11.4	0.4	x	-	x	1.4
1031	Fruit & Veg. Canners & Preservers	x	0.8	3.4	x	1.2	1.5
1040	Dairy Products	2.7	7.7	8.0	3.1	x	5.3
1050	Flour & Breakfast Cereal Prods.	x	0.7	2.9	2.9	-	1.5
1060	Feed	1.1	3.3	2.1	1.6	1.2	2.1
1071	Biscuits	x	0.6	1.0		-	0.5
1072	Bakeries	1.0	0.6	0.5	1.3		0.7
1089	Misc. Food Processors NES	x	4.1	4.3	2.3	2.4	3.2
1093	Breweries	1.2	3.1	2.9	2.6	1.4	2.5
1860	Carpet, Mat and Rug	-	2.5	2.0	x	-	1.3
2310	Hosiery	-	0.5	0.3	-	-	0.2
2392	Other Knitting	x	1.1	x	0.3	-	0.5
2513	Sawmills and Planing Mills	3.7	6.5	3.2	2.5	35.3	10.2
2520	Veneer and Plywood Mills	x	0.7	0.8	1.0	7.4	2.0
2710	Pulp and Paper Mills	26.5	27.3	17.5	6.0	29.6	21.9
2732	Corrugated Box	1.0	1.3	3.6	1.3	1.0	1.9
3550	Ready-Mix Concrete	0.5	1.0	1.5	2.7	0.2	1.2
3651	Petroleum Refining	46.8	27.5	25.3	38.0	14.9	27.9
Per cent of Manufacturing Output[3]							
	Before Exclusions	68.7	35.0	18.9	47.2	66.0	32.0
	After Exclusions	47.4	23.4	11.7	31.4	46.1	21.1

(1) Value of gross output is used. x = confidential and - = no plants in the sample.

(2) The 20 industries sum to 100 per cent in each region. Data after exclusion criteria are used.

(3) The 20-industry sample as a percentage of the whole manufacturing sector. The study sample is measured before data exclusions (i.e., Census data) and after data exclusions (i.e., regression data).

TABLE A3

Comparing the Census Sample and the Regression Sample: Distribution in Percentage of Key Industry Variables by Region Relative to the Canadian Level[1]

Range (Canada = 100)	Atlantic		Quebec		Ontario		Prairies		B.C.	
				- Unit Cost of Production -						
Under 90	5.6	(9.6)	5.0	(3.1)	5.0	(3.1)	5.6	(14.4)	0.0	(9.9)
91 - 98	22.2	(16.6)	20.0	(29.6)	50.0	(28.4)	38.9	(33.5)	33.3	(25.3)
99 - 101	11.1	(16.4)	25.0	(31.5)	10.0	(31.5)	16.7	(13.1)	20.0	(16.2)
102 - 109	50.0	(24.6)	45.0	(30.2)	35.0	(35.1)	22.2	(21.2)	26.7	(23.2)
Over 110	11.1	(32.8)	5.0	(5.6)	0.0	(1.9)	16.6	(17.8)	20.0	(25.4)
				- Average Gross Output by Plant -						
Under 80	72.3	(71.3)	35.0	(32.1)	25.0	(13.0)	61.1	(65.8)	33.3	(71.1)
81 - 90	0.0	(4.9)	5.0	(9.9)	10.0	(6.8)	5.6	(1.3)	0.0	(5.7)
91 - 110	16.7	(9.8)	15.0	(25.9)	35.0	(23.4)	11.1	(15.8)	6.7	(9.1)
111 - 120	0.0	(1.6)	10.0	(7.4)	5.0	(9.3)	0.0	(2.0)	6.7	(2.1)
Over 121	11.0	(12.3)	35.0	(24.7)	25.0	(47.5)	22.3	(15.1)	53.3	(12.0)

(1) The whole Census sample appears in brackets. The sub-sample results are based on actual average data in the sample, not in the Census. Each column sums to 100 per cent.

TABLE A4

Estimated Coefficients of Equation 1 (Nominal Estimates)[1]

Industry	Intercept	V	1/V	W1	W2	DIV	D1	D2	D3	D4
1011	-0.0349	7.3E-5	0.0315	0.0021	-0.0045*	-0.0183**	-0.0302	0.0024	0.0303*	0.0043
	(0.0776)	(1.2E-4)	(0.0767)	(0.0082)	(0.0030)	(0.0091)	(0.0408)	(0.0219)	(0.0199)	(0.0726)
1012	-0.2579**	-0.0003	-0.2253**	0.0117	0.0051*	0.0359	0.2529***	0.0156	-0.0320	-0.0468
	(0.1147)	(0.0006)	(0.1119)	(0.0147)	(0.0036)	(0.0418)	(0.0483)	(0.0248)	(0.0333)	(0.0417)
1020	-0.5712***	0.0050	0.2601***	0.0356	0.0096	0.0194	0.0187	-0.0072	-	0.0139
	(0.2053)	(0.0040)	(0.1084)	(0.0287)	(0.0083)	(0.0214)	(0.0347)	(0.0298)		(0.1510)
1031	-0.4598***	-0.0012***	0.5220***	0.0136	0.0010	0.0264***	0.1010	-0.0302	0.0515	0.1069**
	(0.0950)	(0.0005)	(0.1601)	(0.0145)	(0.0013)	(0.0109)	(0.1043)	(0.0522)	(0.0533)	(0.0574)
1040	-0.2179***	-0.0003	0.0664	-0.0045	0.0099***	0.0211	0.0218	0.0215	-0.0116	-0.0683
	(0.0479)	(0.0003)	(0.0532)	(0.0050)	(0.0027)	(0.0173)	(0.0248)	(0.0231)	(0.0223)	(0.0655)
1050	0.2677	-0.0061***	-0.7472*	0.0125	-0.0116***	-0.1236**	-0.1315**	0.0376	0.0020	-
	(0.3439)	(0.0018)	(0.4982)	(0.0417)	(0.0045)	(0.0636)	(0.0698)	(0.555)	(0.0550)	
1060	-0.0507	-0.0056**	-0.0147	-0.0023	-0.0004	0.0195	0.0006	0.0127	0.0145	0.0943**
	(0.0939)	(0.0032)	(0.2789)	(0.0085)	(0.0023)	(0.0221)	(0.0466)	(0.0314)	(0.0332)	(0.0480)
1071	-0.1975	-0.0022	0.1100*	0.0073	-0.0138	0.0632*	-0.0087	0.0304	0.0524	-
	(0.4591)	(0.0035)	(0.0682)	(0.0455)	(0.0219)	(0.0454)	(0.1213)	(0.0941)	(0.0885)	
1072	-0.5038***	0.0028	0.1143***	0.0133	0.0164**	-0.0171	0.0532	-0.0497	-0.0842***	-0.1854***
	(0.1445)	(0.0025)	(0.0360)	(0.0121)	(0.0082)	(0.0448)	(0.0535)	(0.0448)	(0.0352)	(0.0592)
1089	-0.4746***	-0.0016***	0.1076	0.0461***	-0.0045	-0.0239*	0.0005	-0.0175	-0.1626***	-0.0883
	(0.1219)	(0.0005)	(0.2057)	(0.0139)	(0.0046)	(0.0180)	(0.0997)	(0.0574)	(0.0655)	(0.0866)
1093	-0.8400***	-0.0016	0.8281***	0.0107	-0.0011	-	0.1443**	0.2723***	0.1608***	0.2491***
	(0.2590)	(0.0015)	(0.3760)	(0.0206)	(0.0135)		(0.0718)	(0.0910)	(0.0666)	(0.0308)
1860	0.2562	-0.0035***	0.2642	-0.0245	0.0253	-0.4149	-	0.0585	-0.0442	-
	(0.3441)	(0.0011)	(0.5287)	(0.0600)	(0.0238)	(0.3630)		(0.0638)	(0.0840)	
2310	-0.2338*	-0.0146**	0.0352	0.0090	0.0033	-	-	-0.0158	-	-
	(0.1672)	(0.0077)	(0.0967)	(0.0271)	(0.0056)			(0.0529)		
2392	-0.6185***	0.0060	0.1495*	0.0552**	0.0036	0.0013	-0.0781	-0.0194	-0.0993	-
	(0.2230)	(0.0075)	(0.1112)	(0.0294)	(0.0032)	(0.0128)	(0.1031)	(0.0345)	(0.0935)	
2513	-0.4197***	0.0005	0.0941***	0.0078	0.0007	0.0126	0.0590**	0.0165	0.0384	-0.0012
	(0.0813)	(0.0004)	(0.0310)	(0.0079)	(0.0028)	(0.0229)	(0.0329)	(0.0251)	(0.0561)	(0.0422)
2520	-0.5073***	-0.0003	0.2096**	0.0077	0.0028	0.0825**	-0.0050	0.0694	0.0039	0.1235*
	(0.1589)	(0.0022)	(0.1108)	(0.0182)	(0.0064)	(0.0424)	(0.0835)	(0.0666)	(0.0826)	(0.0963)
2710	-0.5639***	-0.0010***	0.8281***	0.0035	0.0142*	0.0668***	0.0383	-0.0511*	0.0034	-0.1567***
	(0.1754)	(0.0003)	(0.2876)	(0.0133)	(0.0089)	(0.0140)	(0.0471)	(0.0320)	(0.0875)	(0.0498)
2732	-0.2414**	0.0003	0.1530	0.0164*	-0.0069	-0.0007	-0.0227	0.0311*	-0.0347	-0.0344
	(0.1330)	(0.0019)	(0.1298)	(0.0109)	(0.0061)	(0.0421)	(0.0375)	(0.0231)	(0.0274)	(0.0786)
3550	-0.6316***	0.0046*	0.2961***	0.0361***	0.0044	-0.0713	0.1003***	0.0267	-0.0279	-0.0472
	(0.1123)	(0.0030)	(0.0675)	(0.0069)	(0.0044)	(0.0606)	(0.0432)	(0.0319)	(0.0295)	(0.0594)
3651	-0.0872	-3.0E-7	2.8299**	-0.0061	0.0063	-	-0.0171	-0.0660***	-0.0385**	-0.0489*
	(0.1022)	(5.7E-5)	(1.6180)	(0.0109)	(0.0054)		(0.0432)	(0.0259)	(0.0235)	(0.0308)

1. * Significant at the 10% level; ** significant at the 5% level; ***significant at the 1% level. One tail t-test. Standard errors in parenthesis.

TABLE A5

Estimated Coefficients of Equation 2 (Real Estimates)[1]

Industry	Intercept	Q	1/Q	W1	W2	DIV	D1	D2	D3	D4
1011	-0.2501**	-0.0004***	0.5120***	0.0307**	-0.0024	0.0021	-0.1168***	-0.1984***	-0.1173***	-0.0072
	(0.1440)	(0.0001)	(0.0561)	(0.0150)	(0.0048)	(0.0137)	(0.0479)	(0.0510)	(0.0481)	(0.0514)
1012	-0.1077	7.6E-5	-0.0014	-0.0106	0.0004	0.0284	0.4126***	-0.0237	0.0248	0.0640
	(0.1737)	(0.0006)	(0.1792)	(0.0237)	(0.0063)	(0.0476)	(0.0534)	(0.0314)	(0.0560)	(0.0576)
1020	-0.8209***	0.0021	0.1399	0.1292***	0.0177**	0.0396	-0.2553***	-0.1159**	-	-0.5623***
	(0.2324)	(0.0054)	(0.1262)	(0.0379)	(0.0083)	(0.0345)	(0.0696)	(0.0547)		(0.2201)
1031	-0.4615***	-0.0008	0.6318***	0.0324*	-0.0032*	0.0093	0.0095	-0.0667*	-0.0503	0.1367**
	(0.1497)	(0.0008)	(0.1398)	(0.0242)	(0.0020)	(0.0174)	(0.0744)	(0.0462)	(0.0618)	(0.0781)
1040	-0.2966***	-0.0007***	0.2154***	0.0067	0.0014	0.0332**	0.0708*	0.0317	-0.0394	-0.0477
	(0.0933)	(0.0002)	(0.0857)	(0.0108)	(0.0041)	(0.0171)	(0.0443)	(0.0331)	(0.0349)	(0.0922)
1050	-0.3481	-0.0078***	-0.4839	0.0969	-0.0198***	-0.0631	-0.0497	0.0696	-0.0921*	-
	(0.6274)	(0.0030)	(0.8620)	(0.0789)	(0.0070)	(0.0782)	(0.1252)	(0.1097)	(0.0713)	
1060	0.0461	-0.0027	0.0470	-0.0053	-0.0033	-0.0295	0.0688*	0.0368	-0.0750*	0.0355
	(0.1098)	(0.0029)	(0.2428)	(0.0103)	(0.0069)	(0.0344)	(0.0517)	(0.0323)	(0.0531)	(0.0586)
1071	-0.6954*	-0.0098**	0.0830	0.1386***	-0.0202	-0.1301*	0.2034*	0.1097	0.1574*	-
	(0.5344)	(0.0048)	(0.1420)	(0.0515)	(0.0294)	(0.0807)	(0.1383)	(0.1274)	(0.1116)	
1072	-0.5811***	-0.0033	0.2359***	0.0259*	0.0076	-0.0113	0.0085	-0.0870*	-0.1813***	-0.2104**
	(0.1691)	(0.0031)	(0.0403)	(0.0181)	(0.0114)	(0.0733)	(0.0960)	(0.0583)	(0.0699)	(0.0968)
1089	-0.6686***	-0.0014**	0.3904***	0.0718***	-0.0123***	-0.0161	0.1861***	0.0682	-0.1738***	-0.0994**
	(0.1326)	(0.0006)	(0.1656)	(0.0155)	(0.0045)	(0.0184)	(0.0788)	(0.0595)	(0.0611)	(0.0477)
1093	-1.7104***	-0.0010	-1.5232***	0.0696***	0.0140	-	0.2011***	0.2043***	-0.0036	0.2094***
	(0.2557)	(0.0011)	(0.2723)	(0.0170)	(0.0117)		(0.0714)	(0.0673)	(0.0623)	(0.0837)
1860	1.2082**	0.0031	2.7053***	-0.1410**	0.0852***	-1.3359***	-	-0.0497	0.5018***	-
	(0.6112)	(0.0027)	(0.6824)	(0.0836)	0.0364)	(0.4820)		(0.0976)	(0.1325)	
2310	0.3281	0.0085	0.0596	-0.0591	-0.0227***	-	-	-0.2631***	-	-
	(0.3437)	(0.0244)	(0.1147)	(0.0588)	(0.0091)			(0.1106)		
2392	-0.8887***	-0.0082*	0.2501*	0.0839	0.0012	-0.0177	0.1663*	0.1412	0.2610**	-
	(0.3801)	(0.0055)	(0.1560)	(0.0748)	(0.0090)	(0.0330)	(0.1220)	(0.1285)	(0.1517)	
2513	-0.4179***	-0.0002	0.0934***	0.0218**	0.0058**	-0.0066	-0.0203	-0.0423*	-0.1835***	-0.1721***
	(0.0908)	(0.0006)	(0.0405)	(0.0099)	(0.0035)	(0.0269)	(0.0391)	(0.0316)	(0.0630)	(0.0531)
2520	-0.4825**	0.0038	0.3427***	0.0056	-0.0018	0.0049	0.2200*	0.2020***	0.0828	0.2115**
	(0.2397)	(0.0040)	(0.1326)	(0.0236)	(0.0091)	(0.0808)	(0.1432)	(0.0847)	(0.0895)	(0.1159)
2710	-0.5992***	-0.0011***	0.9011***	0.0060	0.0164***	0.0671***	0.0519	-0.0497*	-0.0070	-0.1959***
	(0.1540)	(0.0003)	(0.3369)	(0.0135)	(0.0080)	(0.0164)	(0.0412)	(0.0361)	(0.0833)	(0.0521)
2731	-0.6947***	-0.0062**	0.5367***	0.0577***	-0.0001	0.0809*	-0.0372	0.0781*	-0.1070***	-0.1258
	(0.1975)	(0.0027)	(0.2286)	(0.0183)	(0.0090)	(0.0581)	(0.0581)	(0.0494)	(0.0397)	(0.1251)
3550	-0.7186***	0.0007	0.3655***	0.0286***	0.0025	-0.0291	0.2422***	0.0784**	0.2430***	0.0912
	(0.1433)	(0.0030)	(0.0806)	(0.0111)	(0.0054)	(0.0687)	(0.0464)	(0.0383)	(0.0531)	(0.0775)
3651	0.0048	-7.5E-5***	0.3727	-0.0088	0.0034	-	0.0322**	0.0133	-0.0687***	-0.0376*
	(0.0664)	(3.2E-5)	(1.6190)	(0.0075)	(0.0027)		(0.0186)	(0.0254)	(0.0208)	(0.0239)

1. See Table A4.

TABLE A6

Estimated Coefficients of Equation 3 (Full Estimates)[1]

Industry	Intercept	Q	1/Q	W1	W2	DIV	D1	D2	D3	D4	Pm	Pe
1011	-0.9361***	-0.0002**	0.4676***	0.0228**	-0.0035	0.0081	-0.1137***	-0.1493***	-0.0722**	0.0539	0.5116***	0.1905**
	(0.1884)	(0.0001)	(0.0596)	(0.0117)	(0.0043)	(0.0130)	(0.0489)	(0.0424)	(0.0403)	(0.0626)	(0.0927)	(0.0865)
1012	-0.3985	8.6E-8	-0.0340	-0.0118	0.0014	0.0313)	(0.3388***	-0.0079	0.0448	0.0542	0.3658*	-0.0759
	(0.4114)	(0.0008)	(0.1906)	(0.0224)	(0.0063)	(0.0490)	(0.1186)	(0.0338)	(0.0482)	(0.0587)	(0.2550)	(0.3009)
1020	-1.0035***	0.0013	0.1527	0.1327***	0.0153**	0.0338	-0.1704**	-0.0480	-	-0.4913**	0.1369**	-0.0163
	(0.3395)	(0.0050)	(0.1222)	(0.0404)	(0.0093)	(0.0349)	(0.1009)	(0.0609)		(0.2586)	(0.0716)	(0.1818)
1031	-0.7841***	-0.0016**	0.5282***	0.0301**	-0.0049***	0.0095	-0.0463	-0.0736**	-0.0341	0.1540**	0.1221	0.2566**
	(0.2393)	(0.0009)	(0.1446)	(0.0224)	(0.0018)	(0.0184)	(0.0751)	(0.0416)	(0.0666)	(0.0731)	(0.1489)	(0.1579)
1040	-0.3860**	-0.0007***	0.2051***	0.0074	0.0014	0.0310**	0.0549	0.0308	-0.0404	-0.0479	0.0033	0.0891
	(0.1822)	(0.0002)	(0.0865)	(0.0110)	(0.0041)	(0.0177)	(0.0524)	(0.0326)	(0.0343)	(0.0986)	(0.1064)	(0.1192)
1050	-1.0702***	-0.0072***	-0.2702	0.0450	-0.0022	-0.1118	-0.0670	0.1047	0.1288*	-	0.8895***	0.0255
	(0.3463)	(0.0016)	(0.5807)	(0.0540)	(0.0076)	(0.1375)	(0.0823)	(0.0894)	(0.0798)		(0.3059)	(0.2225)
1060	-1.0127***	0.0020	0.2494	-0.0160***	-0.0002	-0.0075	0.0357	0.0028	0.0567**	0.0300	0.9773***	-
	(0.0505)	(0.0022)	(0.2296)	(0.0067)	(0.0022)	(0.0181)	(0.0301)	(0.0270)	(0.0302)	(0.0568)	(0.1421)	
1071	-0.6942	-0.0098**	0.2750	0.1264*	-0.0031	-0.0304	0.3835	0.1652	-0.0123	-	0.4329	-0.6289
	(2.1124)	(0.0046)	(0.3502)	(0.0814)	(0.0374)	(0.1058)	(0.6033)	(0.1348)	(0.1723)		(1.4466)	(0.5754)
1072	-0.7625***	-0.0040	0.2291***	0.0335**	0.0077	0.0221	-0.0011	-0.1494***	-0.2537***	-0.2753***	0.4890***	-0.3535
	(0.3240)	(0.0034)	(0.0401)	(0.0194)	(0.0116)	(0.0697)	(0.1179)	(0.0554)	(0.0763)	(0.1065)	(0.2009)	(0.2929)
1089	-1.1880***	-0.0011**	0.2633*	0.0729***	-0.0113***	-0.0265**	0.1613***	0.0600	-0.1019**	-0.0776**	0.1713	0.3558**
	(0.2036)	(0.0006)	(0.1759)	(0.0137)	(0.0045)	(0.0160)	(0.0658)	(0.0523)	(0.0473)	(0.0427)	(0.1519)	(0.1348)
1093	-1.9814***	-0.0013*	1.2032***	0.0693***	0.0082	-	0.1478**	0.2112***	-0.0109	0.2148***	0.3180*	0.0735
	(0.2933)	(0.0010)	(0.3066)	(0.0240)	(0.0096)		(0.0681)	(0.0727)	(0.0521)	(0.0846)	(0.2100)	(0.1224)
1860	1.4970*	0.0032	3.0166**	-0.1337*	0.0835**	-1.4324***	-	-0.0371	0.4840***	-	-0.1120	-0.1350
	(1.0008)	(0.0025)	(1.4074)	(0.0983)	(0.0337)	(0.5908)		(0.1154)	(0.1623)		(0.4531)	(0.4214)
2310	-0.3095*	-0.0089	0.1039*	-0.0385	-0.0108*	-	-	-0.2976***	-	-	0.4761***	-
	(0.2371)	(0.0193)	(0.0781)	(0.0381)	(0.0081)			(0.0757)			(0.1683)	
2392	-1.5002*	-0.0052	0.2672**	0.1024	0.0017	-0.0197	0.0047	0.1855*	0.2514**	-	0.1519	0.3100
	(0.9640)	(0.0072)	(0.1560)	(0.0895)	(0.0096)	(0.0381)	(0.3229)	(0.1360)	(0.1477)		(0.1566)	(0.4991)
2513	-0.8034***	-0.0002	-0.0850**	0.0216***	0.0043*	-0.0130	0.0457	-0.0228	-0.0654	-0.1778***	0.3101***	0.0702
	(0.1212)	(0.0005)	(0.0422)	(0.0089)	(0.0032)	(0.0241)	(0.0368)	(0.0288)	(0.0569)	(0.0479)	(0.0347)	(0.0649)
2520	-0.8059***	0.0021	0.3954***	0.0088	-0.0035	0.0198	0.2554**	0.2157***	0.1390*	0.2234**	0.3053***	-
	(0.2439)	(0.0036)	(0.1117)	(0.0220)	(0.0077)	(0.0702)	(0.1192)	(0.0911)	(0.0888)	(0.1114)	(0.1033)	
2710	-0.7986***	-0.0013***	0.9920***	0.0038	0.0175***	0.0708***	0.0906**	-0.0426	0.0262	-0.1363**	0.1995**	-0.0099
	(0.2572)	(0.0003)	(0.3353)	(0.0131)	(0.0072)	(0.0173)	(0.0441)	(0.0443)	(0.0906)	(0.0711)	(0.1164)	(0.1159)
2732	-0.7367**	-0.0062**	0.5212**	0.0567***	0.0005	0.0842*	-0.0343	0.0751*	-0.1151***	-0.1198	0.0904	-0.0461
	(0.3932)	(0.0029)	(0.2335)	(0.0187)	(0.0095)	(0.0581)	(0.0732)	(0.0477)	(0.0393)	(0.1247)	(0.2237)	(0.1206)
3550	-1.1513***	-0.0009	0.2164***	0.0193***	-0.0013	-0.0431	0.1784***	0.0676**	0.0867*	-0.0074	0.5902***	0.0992
	(0.1778)	(0.0027)	(0.0627)	(0.0078)	(0.0043	(0.0638)	(0.0480)	(0.0379)	(0.0589)	(0.0703)	(0.1329)	(0.1482)
3651	-0.7406**	-6.1E-5**	1.7515	-0.0088*	0.0032*	-	0.0474**	0.0033	-0.0500**	-0.0310**	0.7141**	-
	(0.4433)	(3.4E-5)	(1.9232)	(0.0058)	(0.0025)		(0.0256)	(0.0271)	(0.0232)	(0.0179)	(0.4295)	

1. See Table A4.

TABLE A7

Differences Between Full and Nominal Estimates: Hausman Tests[1,2]

Industry	Intercept	Q	1/Q	W1	W2	DIV	D1	D2	D3	D4
1011	-0.9012***	-0.0003***	0.4361	0.0208***	0.0010	0.0264***	-0.0836***	-0.1517***	-0.1024***	0.0496
	(0.1717)	(6.3E-5)	(0.0000)	(0.0083)	(0.0031)	(0.0093	(0.0270)	(0.0363)	(0.0350)	(0.0000)
1012	-0.1406	0.0003	0.1913	-0.0235**	-0.0037	-0.0046	0.0858	-0.0234	0.0767***	0.1011***
	(0.3950)	(0.0005)	(0.1543)	(0.0168)	(0.0051)	(0.0255)	(0.1083)	(0.0230)	(0.0348)	(0.0414)
1020	-0.4323*	-0.0037	-0.1074**	0.0971***	0.0056*	0.0144	-0.1890**	-0.0408	-	-0.5052***
	(0.2704)	(0.0030)	(0.0564)	(0.0284)	(0.0043)	(0.0276)	(0.0947)	(0.0531)		(0.2099)
1031	-0.3243*	-0.0004	0.0062	0.0164	-0.0059***	-0.0169	-0.1473	-0.0435*	-0.0856**	0.0471
	(0.2197)	(0.0008)	(0.0000)	(0.0171)	(0.0012)	(0.0149)	(0.0000)	(0.0317)	(0.0400)	(0.0452)
1040	-0.1681	-0.0004	0.1387**	0.0119	-0.0086***	0.0099***	0.0331	0.0093	-0.0288	0.0204
	(0.1758)	(0.0000)	(0.0682)	(0.0098)	(0.0031)	(0.0039)	(0.0462)	(0.0231)	(0.0261)	(0.0738)
1050	-1.3379***	-0.0010	0.4700*	0.0325	0.0094*	0.0118	0.0645*	0.0670	0.1268**	-
	(0.0403)	(0.0000)	(0.2982)	(0.0342)	(0.0062)	(0.1219)	(0.0435)	(0.0701)	(0.0579)	
1060	-0.9620	0.0076	0.2641	-0.0137	0.0002	-0.0270	0.0351	-0.0099	0.0423	-0.0643**
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.000)	(0.0000)	(0.0000)	(0.0303)
1071	-0.4967	-0.0076**	0.1650	0.1191*	0.0108	-0.0936	0.3921	0.1349	-0.0647	-
	(2.0619)	(0.0031)	(0.3435)	(0.0676)	(0.0303)	(0.0956)	(0.5909)	(0.0966)	(0.1479)	
1072	-0.2587	-0.0068***	0.1148***	0.0202*	-0.0087	0.0392	-0.0543	-0.0997***	-0.1695***	-0.0899
	(0.2900)	(0.0023)	(0.0176)	(0.0152)	(0.0083)	(0.0534)	(0.1050)	(0.0326)	(0.0677)	(0.0885)
1089	-0.7134***	0.0004*	0.1557	0.0268	-0.0068	-0.0026	0.1609	0.0775	0.0607	0.0107
	(0.1631)	(0.0003)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
1093	-1.1414***	0.0003	0.3752	0.0586***	0.0092	-	0.0035	-0.0611	-0.1717	-0.1756*
	(0.1377)	(0.0000)	(0.0000)	(0.0124)	(0.0000)		(0.0000)	(0.0000)	(0.0000)	(0.1206)
1860	1.2408	0.0066**	2.7524**	-0.1092	0.0583**	-1.0175**	-	-0.0956	0.5282*	-
	(0.9398)	(0.0023)	(1.3043)	(0.0778)	(0.0239)	(0.4662)		(0.0960)	(0.1389)	
2310	-0.0758	0.0057	0.0687	-0.0476**	-0.0141**	-	-	-0.2818***	-	-
	(0.1681)	(0.0177)	(0.0000)	(0.0267)	(0.0059)			(0.0542)		
2392	-0.8817	-0.0112	0.1177	0.0472	-0.0018	-0.0210	0.0828	0.2049*	0.3506***	-
	(0.9378)	(0.0000)	(0.1094)	(0.0845)	(0.0090)	(0.0359)	(0.3060)	(0.1315)	(0.1143)	
2513	-0.3837***	-0.0007**	-0.0091	0.0138***	0.0037***	-0.0256***	-0.0134	-0.0393***	-0.1037***	-0.1766***
	(0.0899)	(0.0003)	(0.0285)	(0.0041)	(0.0015)	(0.0077)	(0.0166)	(0.0142)	(0.0097)	(0.0226)
2520	-0.2986*	0.0024	0.1858***	0.0010	-0.0063*	-0.0627	0.2604***	0.1462**	0.1351***	0.0999**
	(0.1850)	(0.0029)	(0.0146)	(0.0124)	(0.0043)	(0.0560)	(0.0851)	(0.0622)	(0.0326)	(0.0561)
2710	-0.2347	-0.0003*	0.1639	0.0003	0.0033	0.0040	0.0523	0.0086	0.0227	0.0205
	(0.1881)	(0.0002)	(0.1724)	(0.0000)	(0.0000)	(0.0102)	(0.0000)	(0.0306)	(0.0235)	(0.0507)
2732	-0.4953*	-0.0065***	0.3682**	0.0402***	0.0074	0.0848**	-0.0115	0.0441	-0.0804***	-0.0853
	(0.3700)	(0.0022)	(0.1941)	(0.0152)	(0.0072)	(0.0401)	(0.0629)	(0.0418)	(0.0282)	(0.0968)
3550	-0.5197***	-0.0055	-0.0797	-0.0168***	-0.0057***	0.0282*	0.0782***	0.0408**	0.1145**	0.0398
	(0.1378)	(0.0000)	(0.0000)	(0.0037)	(0.0010)	(0.0198)	(0.0209)	(0.0205)	(0.0510)	(0.0375)
3651	-0.6534	-6.1E-5*	-1.0783	-0.0027	-0.0030	-	0.0645	0.0693***	-0.0115	0.0180
	(0.4313)	(4.6E-5)	(1.0396)	(0.0000)	(0.0000)		(0.0000)	(0.0081)	(0.0000)	(0.0000)

1. See Table A4.
2. In doing the Hausman test we assume that estimator of equation (1) - $\hat{a}_{i1}$ - is consistent and asymptotically efficient under H_0 (no misspecification) but biased and inconsistent under H_1 (misspecification). Estimator of equation (3) - $\hat{a}_{i3}$ - is assumed to be consistent under H_0 and H_1, but asymptotically inefficient under H_0 (see Hausman (1978) for more details).
3. In theory $V(\hat{a}_{i3}) - V(\hat{a}_{i1}) = V(\hat{q}) \geq 0$ where $\hat{q} = \hat{a}_{i3} - \hat{a}_{i1}$. However, in practice $V(\hat{a}_{i3}) - V(\hat{a}_{i1}) = \hat{V}(\hat{q})$ could be < 0, and we obtained this result in some cases; the standard errors of these negative estimated variances where set to 0.0 in the table.